Deutsche Forschungs- und Versuchsanstalt für Luft- und Raumfahrt EV (DFVLR)
Forschungszentrum Braunschweig
Zentralbibliothek
33 Braunschweig, Flughafen

1
83

Nachrichtentechnik Elektronik

Deutsche Forschungs- und Versuchsanstalt für Luft- und Raumfahrt EV (DFVLR)
Forschungszentrum Braunschweig
Zentralbibliothek
33 Braunschweig, Flughafen
DFVLR 12-342/1983

33. Jahrgang

VEB VERLAG TECHNIK
Träger des Ordens
Banner der Arbeit

Wissenschaftlich-technische Zeitschrift
für die gesamte elektronische Nachrichtentechnik

Herausgeber:
Kammer der Technik
FV Elektrotechnik

Redaktionsbeirat
Nationalpreisträger Prof. Dr. Dr. eh.M.v.Ardenne, Prof.Dr.H.Bernicke, Nationalpreisträger, Prof. Dr. M. Falter, Prof. Dr. sc. techn. P. Fey, Prof. Dr. sc. techn. G. Fritzsche, Nationalpreisträger Prof. em. Dr.-Ing. Dr.-Ing. eh. H. Frühauf, Prof. Dr.-Ing. habil. U. Frühauf, Prof. em. Dr. W. Heinze, Prof. em. Dr.-Ing. eh. W. Kutzsche, Prof. em. Dr. sc. techn. Dr.-Ing. eh. F. H. Lange, Dr. sc. techn. D. Lochmann, Prof. Dr.-Ing. habil. W. Mansfeld, Dr.-Ing. G. Naumann, Dr.rer. oec. Dipl.-Ing. H. Schindler, Prof. Dr.-Ing. habil. P. Vielhauer, Prof. Dr. rer. nat. habil. H. Völz

Inhalt

ISSN 0092-668X

16. Fachkolloquium Informationstechnik 1983
Akustik und Meßtechnik 2. US

Hart, H.
Meßtechnik im Umbruch — Einfluß der Mikroelektronik auf die Meßtechnik 4

Krambeer, H.
Prüfung von analogen Schaltungen mit gesteuerten Quellen 8

Seifart, M.
Zweidraht-Meßdaten-Erfassungssystem mit meßfühlernaher Digitalisierung 10

Brovkov, W.
12-bit-Analog-Digital-Umsetzer mit stufen-paralleler Verarbeitung 12

Kranz, E.-G.
Algorithmen zur digitalen Effektivwertbestimmung 14

Uhlig, G.
Digitale Funktionserzeugung nach dem Tabellenverfahren 16

Püschel, K.
Geräte zur phonetischen Eingabe und Ausgabe — gegenwärtiger Stand und Entwicklungstendenzen 18

Blutner, F.
Psychoakustisches Funktionsmodell und Möglichkeiten einer technischen Systemoptimierung 21

Kordon, U.
Clusterung zur Beurteilung von Analyseverfahren für die Sprachverarbeitung 26

Affa, K.-H.
Verfahren zur dynamischen Sprecheranpassung 28

Weigert, G.
Digitalfilterbank für Sprachanalyse auf einem KRS 4201 30

Haller, K.; Blutner, F.
Multiples Modell zur Beschreibung subjektiver Klangattribute 33

Jaeger, G.
Sprachdaten-Gewinnung nach dem Linear-Predictive-Coding-Verfahren 36

Jugel, A.
Perspektiven der Kommunikation mit Rechnern und Automaten 38

Buchbesprechungen 42

Dissertationen 3.US

Beilage: Jahresinhaltsverzeichnis

Meßtechnik im Umbruch – Einfluß der Mikroelektronik auf die Meßtechnik

H. Hart, KDT, Berlin

Mitteilung aus der Humboldt-Universität zu Berlin, Sektion Elektronik

Deskriptoren: Entwicklung, intelligente Meßgeräte, Meßfühler, Meßtechnik, Meßwertverarbeitung, Mikroelektronik, Prüfautomat, Sensoren

1. Entwicklungsetappen der Meßtechnik

In der Entwicklung der Meßtechnik lassen sich deutlich mehrere Entwicklungsetappen unterscheiden. Unabhängig von anderen möglichen Betrachtungsweisen (z.B. [1]) sind fünf Etappen zu nennen:

In der ersten Etappe dienten Meßverkörperungen zugleich als Meßgeräte. Für die damals wichtigsten Meßgrößen Längen und Mengen waren das vorzugsweise Maßbänder und Lineale sowie Hohlmaße.

In der zweiten Etappe entstanden Meßgeräte, mit deren Hilfe die Meßgröße auf eine andere physikalische Größe abgebildet wurde. So dient beispielsweise in einem Flüssigkeitsthermometer die Länge des Flüssigkeitsfadens als Maß für die Temperatur. Diese Etappe ist auch durch das Entstehen der elektrischen Meßtechnik gekennzeichnet, da elektrische Größen nur auf diese Weise gemessen werden konnten. Dabei war der Erscheinungsort der Meßgröße zugleich Meß- und Ableseort.

Durch die aus der industriellen Meßtechnik stammende Forderung, Meßwerte nicht nur an ihrem Erscheinungsort ablesen zu können, wurden in einer weiteren Entwicklungsetappe neue Lösungswege gefunden. Entweder die Meßgröße wurde an den Ableseort geleitet und dort gemessen oder Messung und Ablesung wurden mit getrennten Geräten realisiert. Ein Beispiel für die erste Möglichkeit ist die Druckfernmessung. Der zweite Weg basiert auf der Nutzung von Signalen zur Übertragung des Abbildungssignals oder des Meßwerts an den Ableseort. Während zunächst auch mechanische Signalträger (Seilzüge, Wellen usw.) benutzt wurden, sind seit einigen Jahrzehnten praktisch nur noch Luft und Elektrizität (pneumatische bzw. elektrische Signale) anzutreffen. In den letzten Jahren ist Licht als eine neue Übertragungsmöglichkeit für Informationen hinzugekommen.

Existierte zuvor für jede Meßgröße noch eine eigene Meßeinrichtung, so ist für die vierte Etappe eine mehr oder weniger umfassende Datenreduktion typisch. Diese Methoden der Meßwertverarbeitung werden einerseits für nur wenige Meßgrößen (oft sogar nur zwei) in kleinen spezialisierten Baugruppen (auch mechanischen [2] und elektrisch-analogen) oder zentral in Meßwerterfassungs- und Meßwertverarbeitungsanlagen realisiert. Mit der stürmischen Entwicklung der Mikroelektronik wurde jedoch die Technik der vom Meßfühler abgesetzten zentralen Meßwertverarbeitung schon bald durch eine dezentrale Verarbeitung der Meßwerte ergänzt [3], die für die gegenwärtige Entwicklungsetappe kennzeichnend ist.

Bei dieser Aufzählung wurde die digitale Meßtechnik nicht als eine gesonderte Entwicklungsetappe betrachtet. Selbstverständlich hat diese Technik eine Reihe von Besonderheiten, aber es gibt von den meßtechnischen Grundsätzen und auch von den benutzten Meßmethoden her keine entscheidenden Unterschiede zur Analogtechnik. Außerdem ist die Entwicklung heute so weit fortgeschritten, daß die digitale Meßtechnik nicht mehr als etwas besonderes behandelt werden sollte, sondern sie muß sich organisch in das Ensemble der herkömmlichen Verfahren einfügen. Im übrigen ist auch die Zuordnung konkreter Geräte zu den analogen oder digitalen keineswegs so einfach und eindeutig, wie es auf den ersten Blick manchmal scheinen mag [4].

2. Stand und derzeitige Möglichkeiten der Meßtechnik

2.1. Entwicklungsstand

Im Gegensatz zu vielen anderen Gebieten der Technik ist in der Meßtechnik das gegenwärtige Entwicklungsniveau dadurch gekennzeichnet, daß jede neue Entwicklungsetappe die meßtechnischen Möglichkeiten erweitert hat, ohne dadurch die früheren Entwicklungen vollständig abzulösen. So werden beispielsweise heute noch mit Lineal, Meßschieber, Winkelmesser usw. weitgehend Maßverkörperungen für die Messung benutzt, wie das vor Jahrtausenden schon der Fall war. Auch das anzeigende Meßgerät, für das Erscheinungsort = Meßort = Ableseort ist, zählt durchaus noch zum Stand der Technik und ist aus der Meßpraxis nicht wegzudenken. Das Weiterleiten der Meßgröße zum Zwecke ihrer Messung an einen entfernten Ort ist zwar nicht mehr sehr verbreitet, die Methode hat aber durch die Lichtleiterpyrometrie (Bild 1) eine neue und durchaus moderne Anwendung gefunden. Die Einzelmessung mit abgesetztem Meßfühler ist in der industriellen Praxis noch absolut üblich, und Meßwerterfassungseinrichtungen mit zentraler Datenreduktion existieren ebenfalls noch, auch in modernen Konfigurationen (Bild 2) [6] [7].

Daneben tauchen andererseits mehr und mehr Meßgeräte der letztgenannten Entwicklungsetappe auf, die allgemein mit dem Ausdruck „intelligente Meßgeräte" treffend charakterisiert werden. Bild 3 zeigt das Schema eines solchen Meßsystems. Der

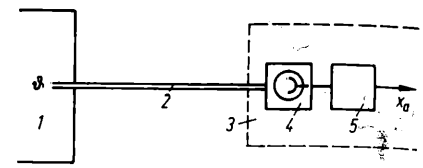


Bild 1. Pyrometrische Temperaturfernmessung durch Weiterleitung der Meßgröße zum Pyrometer [5]
1 Ofen, 2 Lichtleiterkabel (bis zu einigen 100 m Länge), 3 Steuerungseinrichtung, 4 Strahlungsempfänger, 5 Signalwandler, $x_a = f(\vartheta)$

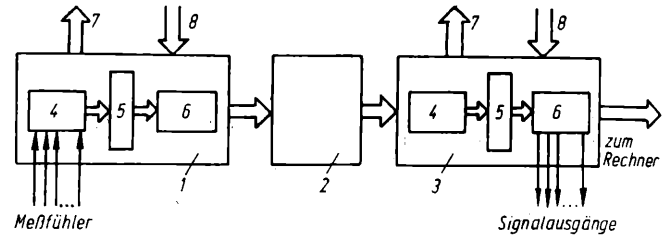


Bild 2. Beispiel einer Meßwertaufbereitung für zentrale Verarbeitung (nach [6])
1 Encoder, 2 Übertragungsmedium, Speicher, 3 Decoder, 4 Datenerfassungsteil, 5 Schnittstelle, 6 Programmteil, 7 Anzeige, 8 Tastatur

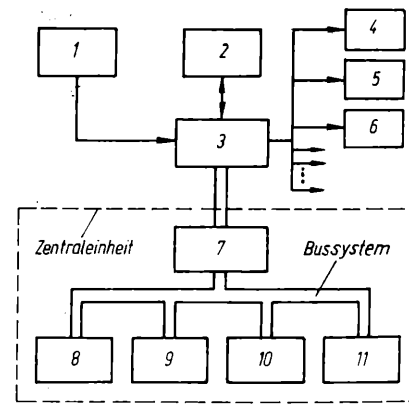


Bild 3. Aufbau eines mikroprozessorgesteuerten Meßsystems (vgl. z. B. [6])
1 Meßfühler, 2 Bedieneinheit, 3 Anpaßeinheit, 4 Anzeige, 5 Drucker, 6 Locher, 7 Eingabe-Ausgabe-Einheit, 8 Mikroprozessor, 9 Zusatzlogik, 10 Programmspeicher, 11 Datenspeicher

Tafel 1. Beispiele für häufig auftretende primäre Abbildungssignale ausgewählter Meßfühler

Größen	Meßfühler Typ	Primäres Abbildungssignal ΔL ($\Delta\alpha$)	ΔU (ΔI)	n
Mechanische Größen	Meßschieber, -schraube	×		
	Differentialtransformator		×	
	Rasterlineal, -scheibe			×
	elastische Kraftaufnehmer DMS-Brücke	×		
	Schwingende Saite			×
	Flüssigkeitsmanometer	×		
	Füllstandsfühler (kapazitiv)		×	
	Füllstandsfühler mechanisch (Schwimmer)	×		
	Füllstandsfühler (Kernstrahlung)		×	×
	Durchflußfühler (rotator.)			×
	Durchflußfühler (induktiv)		×	
	mechanische Waagen	×		
Thermische Größen	Ausdehnungsthermometer	×		
	Thermoelemente		×	
	Widerstandsthermometer		×	
	elektrisches Pyrometer		×	
Strahlungsgrößen	Fotodetektoren		×	
	Bolometer		×	
	Kernstrahlungsdetektoren		×	×

Mikroprozessoreinsatz ist jedoch nur eine Erscheinungsform der verstärkten Nutzung elektronischer und mikroelektronischer Hilfsmittel in der Meßtechnik [9], die darüber hinaus elektrische Funktionsprinzipien erweitern und verbessern sowie viele mechanische Funktionen ersetzen [10].

2.2. Vielzahl der meßtechnisch genutzten Effekte

Bei der Entwicklung von Meßgeräten, insbesondere von Meßfühlern, lassen sich zwei Wege erkennen. Einerseits werden physikalische Effekte in der Forschung entdeckt und nach allen Richtungen auf ihre Ausnutzbarkeit untersucht, wobei sich fast jedesmal herausstellt, daß ein Effekt auch für meßtechnische Zwecke ausgenutzt werden kann. Erscheint diese meßtechnische Nutzung erfolgversprechend, so findet die Entwicklung schließlich in einem verkaufsfähigen Meßgerät ihren Abschluß. Als neueres Beispiel sei auf das Widerstandsrauschen verwiesen, das zur Entwicklung des Rauschthermometers geführt hat, nachdem Untersuchungen ergeben haben, daß dieses Rauschen ausschließlich von der Temperatur abhängt und von anderen Umgebungsparametern praktisch unbeeinflußt bleibt.

Der zweite Weg beginnt mit einem meßtechnischen Problem, das mit den vorhandenen Meßfühlern nicht lösbar ist. Dann werden alle bekannten (und auch vermuteten) Effekte dahingehend überprüft, ob sie für die Lösung dieser Aufgabe in Frage kommen. Ist das der Fall, so entsteht schließlich ebenfalls eine neuartige Lösungsvariante. In der Praxis ist es — wie immer — so, daß die beiden Wege oft kombiniert werden. Eine für die Meßtechnik typische Situation besteht auch darin, daß die Anzahl der meßtechnisch ausgenutzten Effekte ständig zunimmt. Während auf anderen Gebieten normalerweise ältere Lösungen durch neuere abgelöst werden, existieren auf dem Gebiet der Meßfühler vielfach langjährig erprobte und völlig neue Entwicklungen sinnvoll nebeneinander. Nur ganz selten werden Effekte, die einmal für meßtechnische Zwecke genutzt worden sind, wieder fallengelassen, wie es beispielsweise mit der elektrolytischen Abscheidung für die Messung von elektrischer Arbeit [11] geschehen ist.

Die Vielfalt der nutzbaren und auf Grund der unterschiedlichen Meßprobleme und -bedingungen auch benötigten Meßeffekte hat auch eine außerordentliche Vielgestaltigkeit der technischen Realisierungen zur Folge. Die Meßfühler in ihrer Gesamtheit sind kaum noch zu überschauen [8], und die in der übrigen Technik gangbaren Wege zur Vereinheitlichung (und zur Standardisierung) sind hier vielfach nicht begehbar. Eine durchgreifende Veränderung ist auch nicht in Sicht, selbst wenn die moderne Halbleitertechnik in zunehmendem Maße „universelle" Meßfühler anzubieten in der Lage ist (vgl. Abschn. 4.1.).

2.3. Verarbeitung des primären Abbildungssignals

Trotz der großen Vielfalt der genutzten Meßeffekte ist glücklicherweise die Anzahl der primären Abbildungssignale durchaus begrenzt. In mindestens 90% aller Fälle ist das primäre Abbildungssignal eines Meßfühlers eine Längen- oder Spannungsänderung bzw. ein Zählergebnis. Tafel 1 zeigt einige Beispiele dafür, wie die unterschiedlichsten Meßgrößen und die verschiedenartigsten Meßeffekte doch immer wieder auf eines dieser drei Abbildungssignale führen. Diese Tatsache macht es möglich, für die Verarbeitung der Abbildungssignale mehr und mehr vereinheitlichte Baugruppen einzusetzen. Auf diesem Gebiet ist in den vergangenen Jahren schon viel erreicht worden; noch mehr jedoch kann und muß zukünftig in dieser Richtung geschehen. Wenn von Baukastensystemen die Rede ist, so bezieht sich das in der Meßtechnik viel stärker auf diese Baugruppen als auf die Meßfühler.

3. Mikroelektronik in der Meßtechnik

3.1. Verbesserung herkömmlicher Meßgeräte

Während die Mikroelektronik bei den eigentlichen Meßfühlern erst in geringerem Umfang wirksam geworden ist, konnte sie in die übrigen Gebiete der Meßtechnik teilweise sehr schnell eindringen. Natürlich gibt es auch Fühler mit Mikroprozessoren, aber alles, was mit dem primären Abbildungssignal geschieht, gehört eigentlich nicht mehr zur typischen Meßfühlerfunktion, unabhängig davon, ob eine räumliche Trennung besteht oder nicht. Insofern ist auch die Bezeichnung „intelligenter Meßfühler" ein Widerspruch in sich, da sich „Intelligenz" immer auf Verarbeitungsprozesse bezieht.

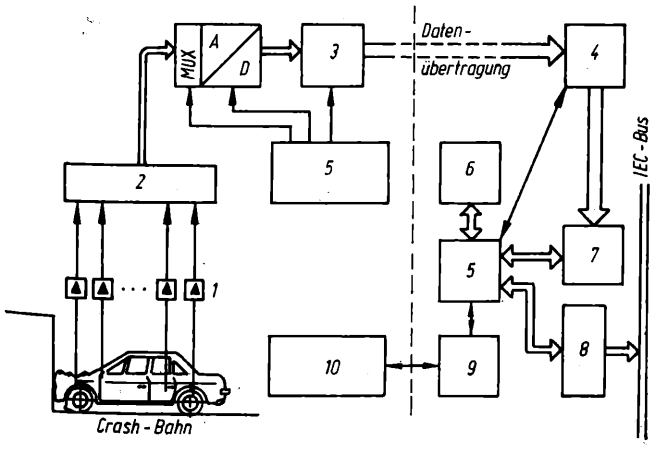
Ist das primäre Abbildungssignal nicht von vornherein eine elektrische Größe (meistens eine Spannung), so ist der erste Verarbeitungsschritt immer die Erzeugung eines elektrischen Signals aus den zunächst auftretenden Längen- und Lichtstromänderungen bzw. Zählergebnissen. Da sich folglich die weitere Signalverarbeitung vollständig auf den elektrischen Bereich beschränkt, hat hier die Mikroelektronik, insbesondere die integrierte Technik, ein außerordentlich nutzbringendes Einsatzgebiet gefunden. So ermöglicht es diese moderne Technik, bessere, kleinere und energiesparende Verstärker, Zähler, Codierer usw. zu schaffen.

Auch die Meßwertausgabe wird mehr und mehr zu einer Domäne der Mikroelektronik bzw. der ihr zugerechneten Optoelektronik. Das gilt nicht nur für Ziffernanzeigen, sondern auch für die analoge Ausgabe, z.B. in Form von Leuchtbalken.

So wichtig derartige Verbesserungen im einzelnen auch sein mögen, so stehen doch bei Betrachtungen über das Eindringen der Mikroelektronik in die Meßtechnik die völlig neuartigen Möglichkeiten berechtigterweise im Mittelpunkt [12]. Die Gebrauchswerterhöhungen der mit Mikroelektronik ausgestatteten Meßmittel sind vor allem durch folgende Aspekte gekennzeichnet:

— Erweiterung der meßtechnischen Möglichkeiten
— Verbesserungen der Meßgenauigkeit
— Erleichterung der Meßwertnutzung
— Erhöhung des Bedienkomforts.

Bild 4. Blockschema der Meßdatenerfassung in Crash-Tests mit Anschlußmöglichkeit für zentrale Auswertung und weitere Systemperipherie über IEC-Bus (nach [15])
1 Meßwandler, 2 Verstärker und Filter, 3 Encoder, 4 Decoder, 5 Steuergerät, 6 Bedienpult, 7 Speicher, 8 Schnittstelle, 9 Steuerrechner, 10 Crash-Bahn-Stellglieder



3.2. Erweiterung der meßtechnischen Möglichkeiten

Eine Erweiterung der meßtechnischen Möglichkeiten besteht zunächst darin, daß sich durch die Automatisierung der Meßvorgänge ihr Ablauf beschleunigen läßt, so daß erhebliche Zeitgewinne erzielbar sind (z.B. [13] [14]) und sogar Vorgänge meßbar werden, die früher nur auf Umwegen erfaßt werden konnten. Automatisierte Meßmittel messen auch dort, wo der Mensch nicht hinkommen kann, z.B. im Weltraum oder auf anderen Planeten. Auch wenn der Einsatz menschlicher Arbeitskraft zu gefährlich (z.B. bei Crash-Versuchen; Bild 4) oder zu unökonomisch (z.B. Dauermessungen im Umweltschutz) wäre, helfen automatische Meßgeräte weiter.

Andererseits gibt es Meßgrößen, die nicht oder nur mit unvertretbarem Aufwand gemessen werden können, da es für sie keinen geeigneten Meßeffekt gibt. Hier ergeben sich neue, bisher längst noch nicht genutzte Möglichkeiten, indem man besser zugängliche Größen mißt und die Aufgabengrößen daraus errechnet. Stoffeigenschaften und -zusammensetzungen sind Größen, die nur in Ausnahmefällen direkt meßbar sind, so daß die Mikroelektronik in der Analysenmeßtechnik ein weites Einsatzfeld findet [16].

Auf Speicheroszillografen, ggf. mit Plottern kombiniert, Logikanalysatoren und andere Meßmittel, die neue Möglichkeiten erschließen, sei hier nur hingewiesen.

3.3. Verbesserungen der Meßgenauigkeit

Die durch die Entwicklung der Mikroelektronik geschaffenen Voraussetzungen für die Erhöhung der Meßgenauigkeit gehen vorzugsweise in vier Richtungen:

— Möglichkeiten zu regelmäßigen automatischen Abgleichen, um die meßtechnischen Eigenschaften der Geräte über längere Zeiten konstanthalten zu können (z.B. [17]).

— Technische Hilfsmittel zur automatischen Korrektion systematischer Fehler entsprechend einer vorliegenden Fehlerkurve. Dazu gehört auch die Linearisierung von nichtlinearen Kennlinien. Sehr wirksam ist auch die kontinuierliche Messung von Einflußgrößen mit automatischer Korrektur der von ihnen verursachten Störeinflüsse.

— Automatisierte Messungen machen Wiederholmessungen (Meßreihen) in einem Umfang möglich, wie es bei Meßdurchführung und -auswertung von Hand nicht denkbar wäre. So sind z.B. Mittelwerte mit sehr kleinen Meßunsicherheiten zu gewinnen.

— Die Zuverlässigkeit der Meßmittel wird erhöht, beispielsweise allein dadurch, daß integrierte Schaltkreise zuverlässiger sind als die gleichen Schaltungen in diskreter Bauweise. Mit der verbesserten Langzeitstabilität wird natürlich auch die Fehlerhäufigkeit von Meßmitteln reduziert [18].

3.4. Erleichterung der Nutzung gemessener Werte

Bekanntlich ist eine Aufgabe, für die Messungen erforderlich sind, in den seltensten Fällen damit gelöst, daß ein Meßwert vorliegt. Abgesehen von den notwendigen Fehlerbetrachtungen sind oft weitere Auswertungen erforderlich. Nicht nur, daß hier die Mikroelektronik erleichternd einspringt, es ist auch möglich, sehr viele Werte sinnvoll zusammenzufassen, also Datenreduktionen vorzunehmen. Die Anzahl derartiger Aufgaben ist unübersehbar. Ob es um die Ermittlung besonders aussagefähiger Größen aus einzelnen verschiedenen Meßgrößen geht (z.B. Ausbeuten, Nutzeffekte, Verbrauchsangaben, Wärmemengen [19]) oder um charakteristische Durchschnittswerte einer Größe (zeitliche oder örtliche Mittelwerte), ob es um Feldverläufe oder um die Bestimmung mehrdimensionaler Größen (z.B. Vektoren) geht, stets ermöglichen intelligente Meßgeräte die Beantwortung der gestellten Frage ohne zusätzliche externe Auswertungen. Auch auf die direkte kontinuierliche Ermittlung statistischer Kenngrößen, z.B. des Verteilungstyps stochastisch schwankender Meßwerte, sei hingewiesen (Bild 5) [20] [21].

Weiterhin ist es möglich, Entscheidungen nach vorgegebenen Gesichtspunkten direkt von der Meßeinrichtung vornehmen zu lassen. Damit sind nicht nur Sortiervorgänge gemeint, sondern auch Entscheidungen über Eingriffe in einen Prozeß, zur Fortsetzung oder zum Abbruch von Messungen (z.B. in Abhängigkeit von der erreichten Standardabweichung), über Auslösung von sicherheitstechnischen Maßnahmen, über die Adressierung von Informationen in Abhängigkeit von Meßergebnissen usw.

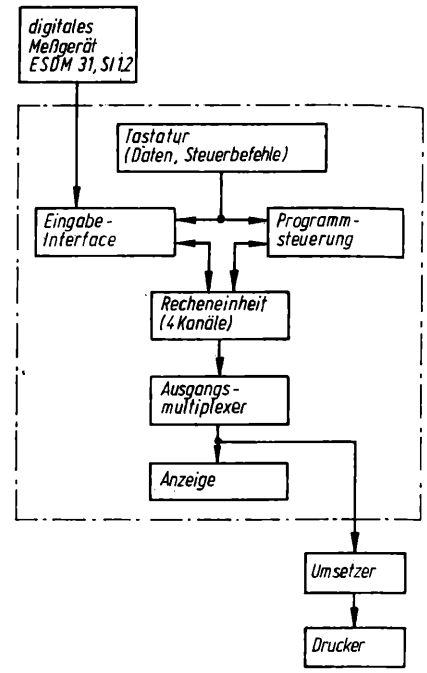


Bild 5. Blockschaltbild eines digitalen Analysators für die statistischen Momente der Verteilung einer Meßwertfolge [21]

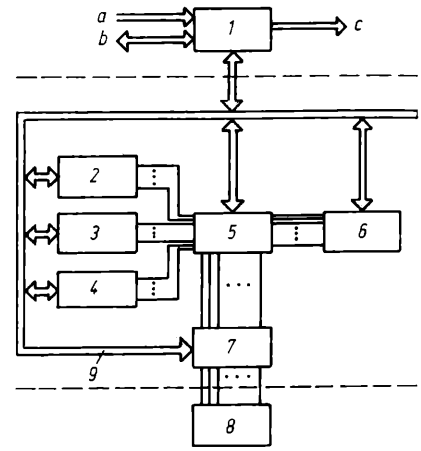


Bild 6. Grundschema eines Meß- und Prüfautomaten
1 Steuereinheit, 2 Signalquellen, 3 Speisequellen, 4 Lasteinheiten, 5 Koppelfeld, 6 Meßsysteme, 7 Interface mit Adapter, 8 Prüfling/Meßobjekt, 9 Bus, *a* Meß-/Prüfprogramm, *b* Operatorinstruktionen, *c* Informationsausgabe und Bedienanforderungen

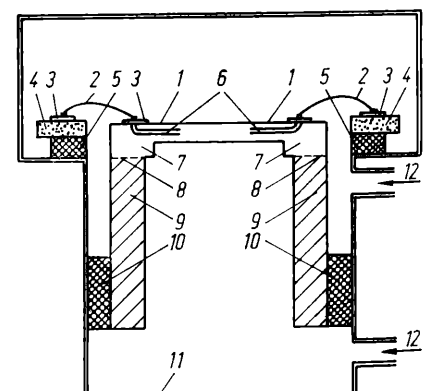


Bild 7. Prinzipieller Aufbau eines piezoresistiven Differenzdruckfühlers (ohne Anschlüsse) (nach [29])
1 ionenimplantierte Piezowiderstände in Brückenschaltung, 2 Bonddrähte, 3 Goldauflagen, 4 Keramikträger, 5 Epoxidharz, 6 interne Zuleitungen, 7 Siliziumchip mit eingeätzter kreisförmiger Druckmembran, 8 molekulare Bindung, 9 Rohr, 10 Silikondichtung, 11 Gehäuse, 12 Druckanschlüsse

Schließlich werden Meßergebnisse auch dadurch überschaubarer, daß man sie grafisch darstellt. Hat man die Meßwerte bei der Messung im Speicher des eingebauten Mikrorechners, so ist es nur noch ein kleiner Schritt bis zu einer optimalen Form der Ergebnisausgabe, also auch zu einer Ausgabe in Kurvenform mit einem Plotter.

3.5. Erhöhung des Bedienkomforts

In der Meßtechnik ist die Vereinfachung der Meßgerätebedienung nicht allein eine Frage der Bequemlichkeit bzw. der Produktivitätserhöhung, sondern sie dient auch der Vermeidung von sub-

jektiven Fehlern, womit höhere Meßgenauigkeiten erreicht und Meßgeräte vor Beschädigungen geschützt werden. Das wird deutlich, wenn man betrachtet, durch welche Einzelmaßnahmen der Bedienkomfort von Meßgeräten mit mikroelektronischen Mitteln erhöht wird. Intelligente Meßgeräte wählen z.B. selbständig den optimalen Meßbereich, achten bei elektrischen Messungen auf richtige Polung, steuern automatisch den Meßablauf, wählen bei laufenden Messungen die günstigste Meßfrequenz in Abhängigkeit von der Änderungsgeschwindigkeit der Meßgröße, melden aufgetretene Störungen usw.

Besonders eindrucksvoll sind die Erfolge der Mikroelektronik bei der Entwicklung der rechnergestützten Prüftechnik (Bild 6) [21] bis [25]. Dieses Einsatzgebiet der Meßtechnik zeigt außerdem sehr anschaulich, wie die Meßtechnik nicht nur von der Mikroelektronik profitiert, sondern wie auch die Entwicklung der Mikroelektronik auf Fortschritte in der Meßtechnik angewiesen ist, da beispielsweise die Kleinheit der Meßobjekte (z.B. Chips von integrierten Schaltkreisen) und die Vielzahl der erforderlichen Einzelmessungen (z.B. kombinatorische Schaltkreise) völlig neue und sehr anspruchsvolle Forderungen an die Meßtechnik stellt [21].

4. Entwicklungstendenzen

4.1. Meßfühler [7] [26] [27] [37]

Wenn auch bereits darauf hingewiesen wurde (s. Abschn. 2.3.), daß auf dem Gebiet der Meßfühler die Auswirkungen der Mikroelektronik weniger augenfällig sind, so gibt es doch einige Entwicklungstendenzen, die zumindest von den im Zusammenhang mit der Mikroelektronik entstandenen technologischen Verfahren profitieren [10] [18] [28]. Als Beispiel werden die Halbleiterfühler, z.B. Halbleiterdehnmeßstreifen und piezoresistive Aufnehmer (Bild 7), genannt. Neu sind auch Silizium-Widerstandsthermometer [30] sowie Sintertechnologien für Gasanalysen-Meßfühler und Dünnfilmtechnologien für CO_2- und Feuchtefühler [31]. Dabei rücken neben dem Silizium auch Mehrstoffhalbleiter (GaAs, TaNi u.a.) ins Blickfeld [37].

Eine andere Entwicklung dagegen läßt sich erst in den Umrissen erkennen, und zwar die echte Einbeziehung von Meßfühlern in integrierte Schaltungen. Hier könnte man beispielsweise an temperaturabhängige Bauelemente denken, die benutzt werden, um Temperatureinflüsse auf die Funktion von bestimmten Schaltungen zu kompensieren. (Natürlich ist es fraglich, ob in diesen Fällen noch von Meßfühlern gesprochen werden kann.)

Da die Optoelektronik als Teilgebiet der Mikroelektronik betrachtet wird, sei auch auf die Meßfühler für nichtoptische Größen mit optischen Ausgangssignalen hingewiesen [5] [32]. Bei ihnen besteht die Zielstellung darin, sie lichtleiterkompatibel zu gestalten (vgl. Abschn. 4.2.), um in der prozeßnahen Ebene ohne elektrische Energie auszukommen. Deshalb wird ein über Lichtleiter zugeführter Lichtstrom durch die Meßgröße moduliert und die Lichtstromänderung als Meßsignal ebenfalls über Lichtleiter zum eigentlichen Meßgerät übertragen. Wegen der Störanfälligkeit amplitudenmodulierter Signale wird bei der weiteren Entwicklung das Augenmerk auf die Gewinnung diskreter Signale gerichtet werden müssen [32]. In dem Maße, wie die Optoelektronik sich zunehmend der integrierten Optik zuwendet [33], wird man auch Meßfühler, die heute als Faserfühler noch in diskreter Technik hergestellt werden [5], in integrierter Technik produzieren, die eine neue Qualität auf dem Meßfühlergebiet sind. Auf andere Neuentwicklungen, wie den Wiegand-Sensor [35], fluidische Sensoren usw. [34], wird nicht eingegangen, da sie kein Ergebnis der Mikroelektronikentwicklung sind.

4.2. Verbesserungen bei der Meßsignalweiterleitung

Auch wenn elektrische Spannungen bzw. elektrische Ströme für die Übertragung von Meßinformationen noch vor einiger Zeit als optimal galten, so sind doch auf diesem Gebiet die Weiterentwicklungen nicht zu übersehen. Mit der Digitalisierung von Meßsignalen, insbesondere der dezentralen direkt am Meßfühler, ist es möglich, Übertragungsfehler nahezu völlig auszuschließen. Im Vergleich zur analogen Übertragung sind dabei Schwankungen des Leitungswiderstands von geringerem Einfluß auf die Meßgenauigkeit, und durch Verwendung fehlererkennender oder sogar -korrigierender Codes wird die Übertragungssicherheit weiter erhöht.

Das niedrige Energieniveau TTL-kompatibler Signale ermöglicht einen leichten Explosionsschutz, und auch der Energieverbrauch der mikroelektronikbestückten Meßgeräte ist beträchtlich gesunken. Damit wird nicht nur die Energieversorgung der normalen Meßgeräte erleichtert (z.B. kleinere Stromversorgungseinheiten), sondern es werden auch Fernmessungen mit leitungsloser Signalübertragung von entfernten Meßstellen möglich, für die Sonnenenergie als Energiequelle genutzt werden kann. Das gilt nicht nur für die Raumfahrt, sondern auch für Meßbojen auf hoher See oder Meßstellen in unzugänglichen Gebirgsregionen.

Neben anderen elektrischen Signalformen ist auch ein neuer Signalträger meßtechnisch interessant geworden: das Licht. Die zunächst für Fernmeldezwecke entwickelte Lichtleitertechnik hat auch in Signalübertragungsnetzen für Meß- und Automatisierungszwecke so viele Vorteile [5] [32], daß sie sich für bestimmte Anwendungsgebiete, beispielsweise in elektrischen Energienetzen, sicherlich durchsetzen wird. Die hohe Übertragungskapazität macht aber diesen Signalträger auch für andere Zwecke, gegebenenfalls sogar innerhalb von Meßsystemen, interessant.

4.3. Intelligente Meßgeräte

Der Einsatz von Mikroprozessoren, d.h. von „dezentraler Intelligenz", in den verschiedensten Meßmitteln steht erst am Anfang. Die als Beispiel aufzufassenden Ausführungen über die neuen meßtechnischen Möglichkeiten im Abschn. 3. lassen den endgültigen Umfang, den mikroelektronische Hilfsmittel in der Meßtechnik einmal einnehmen werden, kaum ahnen. Sicherlich wird die neue Technik die Meßtechnik nicht vollständig ändern, sondern bewährte „klassische" Lösungen werden weiterhin ihre Daseinsberechtigung behalten; das weitere Vordringen der Mikroelektronik ist nicht aufzuhalten. Wurde noch vor Jahren der Meßschieber mit elektronischer Meßsignalverarbeitung und digitaler Anzeige z.T. als eine Art Spielerei abgetan, so zeigt die derzeitige internationale Marktsituation jedoch, daß er dank der erleichterten Ablesbarkeit bei voller Ausnutzung des Auflösungsvermögens vielerorts bereits eindeutig bevorzugt wird. Ähnlich ist es in der Wägetechnik [6] und auf vielen anderen meßtechnischen Gebieten.

In diesem Zusammenhang sei auch darauf hingewiesen, daß das staatliche Meßwesen von diesen Entwicklungen nicht unbetroffen bleibt [13]. Es profitiert einerseits von den verbesserten Voraussetzungen für die Darstellung und Weitergabe der Einheiten, muß aber auch gestiegenen und z.T. völlig neuartigen Forderungen an die Meßmittelprüfung, speziell die Eichung, gerecht werden. Die Einbeziehung von Mikrorechnern und damit auch von Software in Meßmittel stellt diejenigen, die für die Richtighaltung von Meßmitteln verantwortlich sind, vor völlig neuartige Probleme. Auch hier handelt es sich nicht um quantitative Veränderungen, sondern um eine neue Qualität.

4.4. Ökonomie und Optimierung

Ist die Ermittlung des Nutzens, den die Meßtechnik erbringt, an und für sich schon ein weitgehend ungelöstes Problem (vgl. z.B. [36]), so erscheint diese Frage im Zusammenhang mit dem Eindringen der Mikroelektronik in einem völlig neuen Licht [13]. Es sind z.T. die erzielbaren Vorteile ganz offensichtlich, wenn beispielsweise bisher Unmeßbares jetzt meßbar wird (vgl. Abschn. 3.2.), teilweise gehen aber die Ansichten auch weit auseinander.

Die heikle Frage des ökonomischen Nutzens von Genauigkeitserhöhungen (vgl. Abschn. 3.3.) sei hier ausgeklammert, aber auch wie hoch der Nutzen von erhöhtem Bedienkomfort (vgl. Abschn. 3.5.) einzusetzen ist, dürfte sehr von den konkreten Bedingungen abhängen. Dazu kommt, daß die stürmische Entwicklung der Mikroelektronik auch die Preise ständig in Bewegung hält, so daß eine Maßnahme, die heute noch unökonomisch erscheint, morgen schon zweckmäßig sein kann.

Die Entscheidung sowohl für die Entwicklung intelligenter Meßgeräte und anderer mikroelektronikgestützter Meßmittel als auch für deren Einsatz ist ein Optimierungsproblem, dessen Eingangsgrößen ständigen Veränderungen unterliegen. Sachkenntnis, Verantwortungsbewußtsein, prognostisches Denken und Risikobereitschaft sind also nötig, um hier den richtigen Weg zu

finden. Dabei kann im Rahmen einer derartigen Darstellung auf die Problematik nur hingewiesen werden, da das Problem von Fall zu Fall zu verschieden ist, als daß allgemein gehaltene Betrachtungen weiterhelfen könnten.

NaA 8981

Literatur

[1] *Rohr, H.:* Etappen zur MSR-Geräteentwicklung zwischen 1900 und 1965. messen steuern regeln 24 (1981) 9, S. 528 – 534

[2] *Huthloff, E.:* Eichpflichtige Gasmengenmessung mit Wirbelgaszähler und elektrischer Umwertung. Techn. Messen 47 (1980) 4, S. 125 – 130

[3] *Seifart, M.; Schommartz, G.; Frühauf, U.:* Stand und Perspektiven der elektronischen Meßwerterfassungstechnik zwischen Sensor und Mikrorechner. Jahrestagung der WGMA, Jena, 5. und 6. November 1981, Tagungsband S. 45 – 85

[4] *Hart, H.:* Einführung in die Meßtechnik, 3. Aufl. Berlin: VEB Verlag Technik 1980

[5] *Hart, H.; Pertel, R.:* Meßfühler für nichtoptische Meßgrößen auf optischen Prinzipien und mit optischen Ausgangssignalen. Feingerätetechnik 32 (1983) (in Vorbereitung)

[6] *Voigt, D.:* Meßdatenerfassung durch dezentrale Intelligenz. Elektronik 31 (1982) 9, S. 61 – 67

[7] *Obermeier, E.; Reichel, H.:* Meßwerterfassungssysteme und Sensorprinzipien. Elektronik 28 (1979) 26, S. 23 – 29

[8] *Mészáros, L.; Bértenyi, A.:* Mikroprozessoren in automatischen Wägeeinrichtungen. Feingerätetechnik 27 (1978) 5, S. 199 u. 200

[9] *Böhme, L.:* Feingeräte unter dem Einfluß der Mikroelektronik. Feingerätetechnik 27 (1978) 5, S. 194 – 198

[10] *Busch, R.:* Unter die Lupe genommen: Sensoren. m + pr/automatik 17 (1981) 9, S. 546 – 550 u. 553

[11] *Pflier, P. M.:* Elektrizitätszähler. Berlin Göttingen Heidelberg: Springer-Verlag 1954

[12] *Naumann, G.:* Mikrorechneranwendung in der elektronischen Meßtechnik. Nachrichtentechnik · Elektronik 32 (1982) 4, S. 136 – 139

[13] *Hart, H.; Schmidt, K.:* Meßautomatisierung und Metrologie. Jahrestagung der WGMA, Leipzig, 15. und 16. November 1979, Tagungsband S. 104 bis 124

[14] *Thiemer, R.; Kunz, M.:* Automatische Prüfung von Präzisionsteilkreisen mit On-line-Rechnereinsatz. messen steuern regeln 22 (1979) 10, S. 582 – 585

[15] *Kramer, M.:* Integrierte Datenerfassung und -verarbeitung im Fahrzeugsicherheitsversuch. m + pr/automatik 15 (1979) 10, S. 785 – 790

[16] *Gatzmanga, H.; Langhoff, N.:* Stand und Perspektiven der rechnergestützten Analysenmeßtechnik. Jahrestagung der WGMA, Jena, 5. und 6. November 1981, Tagungsband S. 86 – 117

[17] *Wagner, F. E.:* Selbstkalibrierende Meßsysteme. m+pr/automatik 17 (1981) 10, S. 666 – 668, 671 u. 672; 11, S. 757 – 761 und 12, S. 834, 837 – 839

[18] *Kästel, W.; Wörner, D.:* Stand und Entwicklungstendenzen der Meßtechnik. Qualität und Zuverlässigkeit 26 (1981) 11, S. 348 – 351

[19] *Kluge, H.:* Wärmeleistungs- und Wärmemengenmessung – Beispiele für mikrorechnergekoppelte Meßsysteme. messen steuern regeln 24 (1981) 3, S. 126 – 129

[20] *Hart, H.; Rothe, A.:* Statistische Verarbeitung von Prüfdaten im On-line-Betrieb. 3. Int. Fachtagung „Meßinformationssysteme". Karl-Marx-Stadt, 16. bis 18. Oktober 1979, Tagungsband S. 47 – 51

[21] *Hart, H.:* Prüfung von LSI-Schaltkreisen. Probleme der Festkörperelektronik, Bd. 12: Mikroprozessoren. Berlin: VEB Verlag Technik 1982

[22] *Lorenz, E.:* Rechnergesteuerte Automaten für die Prüf- und Check-out-Technik. Techn. Messen 46 (1979) 9, S. 315 – 321

[23] *Hascher, W.:* Im Blickpunkt: Automatische Testsysteme. Elektronik 29 (1980) 5, S. 43 – 49 und S. 51 – 52

[24] *Frühauf, U.; Geiger, L.:* System- und Programmstruktur von Meß- und Prüfautomaten. Nachrichtentechnik · Elektronik 29 (1979) 1, S. 4 – 6

[25] *Kämpfe, G.:* Einsatz mikroprozessorgesteuerter Prüfautomaten bei der Fertigung von Geräten der Übertragungstechnik. Qualität und Zuverlässigkeit 26 (1981) 10, S. 306 – 310

[26] *Hofmann, D.:* Neue Entwicklungstendenzen der Sensortechnik. messen steuern regeln 24 (1981) 12, S. 662 – 667

[27] *Hofmann, D.; Richter, W.:* Stand und Perspektiven der Sensortechnik. Jahrestagung der WGMA, Jena, 5. und 6. November 1981, Tagungsband S. 2 – 28

[28] *Gutgesell, H.:* Integrierte Druckwandler und ihre Anwendung. Elektronik 30 (1981) 1, S. 73 – 76

[29] *Hencke, H.:* Piezoresistive Druckaufnehmer. Elektronik 30 (1981) 8, S. 59 u. 60

[30] *Reuber, C.:* Sensoren – Sinnesorgane der Mikroelektronik. Teil 1. Funk-Technik 37 (1982) 4, S. 132 – 134

[31] *Hofmann, D.:* Anwendung der Mikrorechner in der Meßtechnik (Tagungsbericht) messen steuern regeln 25 (1982) 2, S. 105 u. 106

[32] *Hart, H.:* Zur Entwicklung von Sensoren mit optischem Ausgangssignal für verfahrenstechnische Meßgrößen. Vortrag auf dem 27. IWK, Ilmenau, 25. bis 29. Oktober 1982, Berichtsband 4, S. 43 – 46

[33] *Bischoff, B.:* Mikrostrukturen für Funktionselemente der integrierten Optik. Feingerätetechnik 29 (1980) 6, S. 266 – 269

[34] *Bethe, K.; Meyer-Ebrecht, D.:* Sensoren für die Konsumelektronik. Elektronik 29 (1980) 10, S. 41 – 46

[35] *Bergmann, H.:* Der Wiegand-Sensor. radio fernsehen elektronik 31 (1982) 4, S. 230

[36] *Hart, H.:* Über die Ermittlung des ökonomischen Nutzens bei der Dickenmessung mit Hilfe von Kernstrahlung. Zellstoff und Papier 13 (1964) 1, S. 13 – 17

[37] Sensoren-Technologie und Anwendung. NTG-Fachberichte Bd. 79, Berlin Offenbach: VDE-Verlag GmbH 1982, S. 146 – 151

Prof. Dr. rer. nat. habil. *Hans Hart*, Humboldt-Universität zu Berlin, Sektion Elektronik, 1136 Berlin, Hans-Loch-Str. 349

Prüfung von analogen Schaltungen mit gesteuerten Quellen

H. Krambeer, KDT, Wismar

Mitteilung aus der Ingenieurhochschule Wismar, Sektion Technologie der Elektrotechnik/Elektronik

Deskriptoren: Prüfung, Fehlerlokalisierung, analoge Schaltungen, gesteuerte Elemente, Knotenspannungsanalyse

In früheren Beiträgen in dieser Zeitschrift [1] [2] wurde ein Verfahren zur Fehlerortung in linearen analogen Netzwerken beschrieben. Zwischenzeitlich wurden auch in [3] bis [5] ähnliche Untersuchungen mitgeteilt. Nachfolgend wird ergänzend dargelegt, welche Besonderheiten sich bei Vorhandensein gesteuerter Quellen ergeben. Berücksichtigt wird weiterhin der idealisierte Operationsverstärker (Nullor) als Netzwerkelement [6].

1. Grundlagen des Verfahrens

Der Ausgangspunkt zur Erkennung und der Lokalisierung von Fehlern in einer analogen Schaltung ist deren Beschreibung durch die indefinite Mehrtoradmittanzmatrix $(\boldsymbol{Y})$, die den Zusammenhang zwischen den Einströmungen $(\boldsymbol{I}_0)$ an den Knoten und den meßbaren Knotenpotentialen $(\boldsymbol{U})$ herstellt:

$$(\boldsymbol{I}_0) = (\boldsymbol{Y})\,(\boldsymbol{U}). \tag{1}$$

Es wird ohne Beschränkung der Allgemeinheit vorausgesetzt, daß die Schaltelemente durch Zweipolersatznetzwerke im einfachen Fall durch einen Zweipol beschrieben werden können; dann ist das Ersatzschaltbild der Schaltung ebenfalls ein Zweipolnetzwerk. Im Fall eines gemeinsamen Bezugspotentials für alle Knotenpotentiale sowie Zugänglichkeit aller Knoten für Messungen sind die Elemente $y_{\mu\nu}$ $(\nu \neq \mu)$ von $(\boldsymbol{Y})$ bis auf das Vorzeichen die Zweigadmittanzen $Y_{\mu\nu}$:

$$y_{\mu\nu} = -\,Y_{\mu\nu}. \tag{2}$$

Für passive Zweige gilt $y_{\mu\nu} = y_{\nu\mu}$. Ein Fehler $\Delta Y_{\mu\nu}$ eines passiven Zweipols $Y_{\mu\nu}$ zwischen den Knoten μ und ν verändert genau vier Elemente von $(\boldsymbol{Y})$; nämlich $y_{\nu\nu}$, $y_{\nu\mu}$, $y_{\mu\nu}$ und $y_{\mu\mu}$, da zu diesem Zweipol die folgende Teilmatrix $(\boldsymbol{Y})_{\mu\nu}$ als Summand von $(\boldsymbol{Y})$ gehört:

$$(\boldsymbol{Y})_{\mu\nu} = \begin{matrix} \\ \mu \\ \nu \end{matrix} \begin{matrix} \begin{matrix} \mu & \nu \end{matrix} \\ \begin{bmatrix} Y_{\mu\nu} & -Y_{\mu\nu} \\ -Y_{\mu\nu} & Y_{\mu\nu} \end{bmatrix} \end{matrix}. \tag{3}$$

Man hat also für ein Netzwerk mit einem solchen Einzelfehler

$$(\boldsymbol{I}_0) = \begin{bmatrix} y_{11} & \cdots & y_{1n} \\ \cdot & & \cdot \\ \cdot & y_{\mu\mu} + \Delta Y_{\mu\nu}\, y_{\mu\nu} - \Delta Y_{\mu\nu} & \cdot \\ \cdot & y_{\nu\mu} - \Delta Y_{\mu\nu}\, y_{\nu\nu} + \Delta Y_{\mu\nu} & \cdot \\ y_{n1} & \cdots & y_{nn} \end{bmatrix} \begin{bmatrix} U_1 \\ \cdot \\ \cdot \\ \cdot \\ U_n \end{bmatrix}$$

n Anzahl der Knoten des Netzwerks. Daraus wird durch Umstellung

$$\begin{matrix} \\ \mu \\ \nu \end{matrix} \begin{bmatrix} 0 \\ \vdots \\ -\Delta Y_{\mu\nu}\,(U_\mu - U_\nu) \\ +\Delta Y_{\mu\nu}\,(U_\mu - U_\nu) \\ \vdots \\ 0 \end{bmatrix} = (\boldsymbol{Y})\,(\boldsymbol{U}) - (\boldsymbol{I}_0). \tag{4}$$

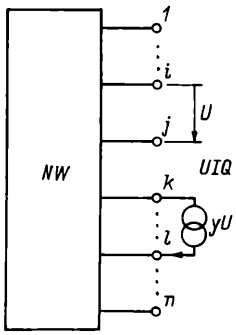


Bild 1. Netzwerk mit spannungsgesteuerter Stromquelle

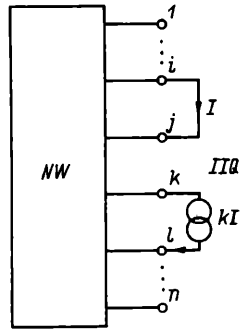


Bild 2. Netzwerk mit stromgesteuerter Stromquelle

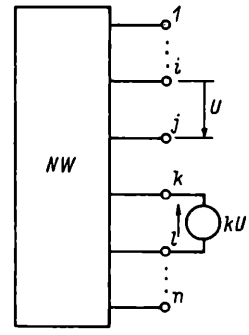


Bild 3. Netzwerk mit spannungsgesteuerter Spannungsquelle

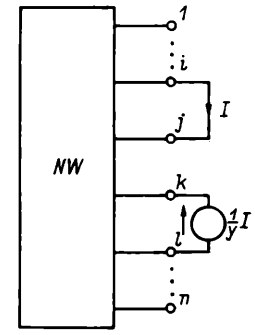


Bild 4. Netzwerk mit stromgesteuerter Spannungsquelle

Die Größen auf der rechten Seite der Gl. (4) sind für die Fehlerortung als bekannt anzusehen: $(\boldsymbol{Y})$ ist die Matrixbeschreibung der fehlerfreien Schaltung, $(\boldsymbol{I}_0)$ enthält die Einströmungen an den Knoten (Ströme der Generatoren), und die Knotenpotentiale $(\boldsymbol{U})$ werden an der zu prüfenden Schaltung gemessen. Die linke Seite von Gl. (4) ist der zu einem Einzelfehler gehörende Fehlerstromvektor [2]. Die von Null verschiedenen Elemente des Fehlerstromvektors geben die Lage des Fehlers an, selbstverständlich ist auch die Größe $\Delta Y_{\mu\nu}$ des Fehlers daraus berechenbar. In prinzipiell gleicher Weise lassen sich so auch Fehler in einem mehrpoligen Schaltelement orten, wenn die Anschlußklemmen dieses Elements durch die äußere Beschaltung nicht direkt verbunden sind.

Die Fehlerlokalisierung bei Auftreten von Mehrfachfehlern ist in [2] erläutert. Es sei besonders darauf hingewiesen, daß für den Fall, daß die fehlerhaften Elemente Maschen im Netzwerk bilden, das in [1] beschriebene Verfahren der Parameterbestimmung herangezogen oder die direkte Messung von Übertragungsadmittanzen notwendig wird (In Circuit-Test).

Das Vorhandensein gesteuerter Quellen hat Auswirkungen auf die Lokalisierbarkeit von Fehlern im übrigen Netzwerk, außerdem ist festzustellen, wie Fehler der Steuerparameter erkannt und die Steuerparameter aus Messungen an der Schaltung bestimmt werden können.

2. Netzwerk mit spannungsgesteuerter Stromquelle (UIQ)

Die für die im Bild 1 dargestellte Zusammenschaltung resultierende Admittanzmatrix ist:

$$(\boldsymbol{Y}) = (\boldsymbol{Y})_{NW} + \begin{matrix} \\ k \\ l \end{matrix} \begin{matrix} \begin{matrix} i & j \end{matrix} \\ \begin{bmatrix} y & -y \\ -y & y \end{bmatrix} \end{matrix}. \tag{5}$$

Ein Fehler des Steuerleitwerts y wirkt sich also in den Elementen k und l des Fehlerstromvektors aus, die gleiche Wirkung hätte ein Fehler im Zweig Y_{kl}. Wenn also ein solcher Zweig in der zu prüfenden Schaltung vorhanden ist, muß, wenn ein Fehler zwischen den Knoten k und l auftritt, wenigstens einer der mit y verknüpften Übertragungsleitwerte gemessen oder nach [1] ermittelt werden. Die Fehlerfreiheit der übrigen Elemente kann in der im Abschn. 1. beschriebenen Weise festgestellt werden.

3. Netzwerk mit stromgesteuerter Stromquelle (IIQ)

Die Beschaltung eines Netzwerks NW mit einer stromgesteuerten Stromquelle entsprechend Bild 2 wird durch folgendes Gleichungssystem beschrieben:

$$\begin{bmatrix} I_1 \\ \vdots \\ I_i + I_j \\ I_k + I_l \\ \vdots \\ I_n \end{bmatrix} = (\boldsymbol{Y})_a \begin{bmatrix} U_1 \\ \vdots \\ U_i \\ \vdots \\ U_n \end{bmatrix}. \tag{6}$$

Dabei wurden der Steuerstrom- und auch der Steuerparameter eliminiert. Gleichzeitig gilt $U_i = U_j$. $(\boldsymbol{Y})_a$ geht also aus $(\boldsymbol{Y})_{NW}$ durch Addition der i-ten und j-ten Zeile, der k-ten und der l-ten Zeile sowie der i-ten und j-ten Spalte hervor. Damit sind mit Hilfe der Fehlerstrommethode die Fehler

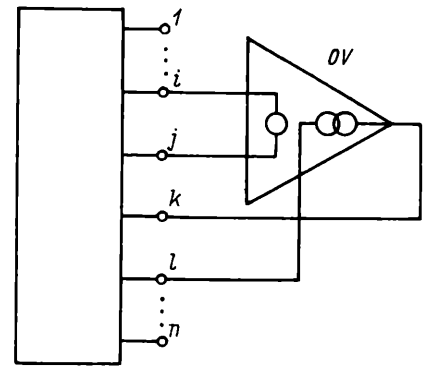


Bild 5. Netzwerk mit idealem Operationsverstärker (Nullor)

$\Delta Y_{\nu i}$ und $\Delta Y_{\nu j}$ sowie die Fehler

$\Delta Y_{\nu k}$ und $\Delta Y_{\nu l}$ $(\nu = 1 \ldots n)$

nicht getrennt zu erkennen. Insbesondere hat ein Fehler ΔY_{kl} keine Wirkung im System (6). Ein Zweig Y_{ij} in der Schaltung ist ohne Bedeutung, da er durch den steuernden Zweig kurzgeschlossen wird.

Weiterhin gelten die folgenden Gleichungen:

$$I_k - \sum_{\substack{\nu=1 \\ \neq k,l}}^{n} y_{k\nu} U = k I_i - \sum_{\substack{\nu=1 \\ \neq k,l}}^{n} k y_{i\nu} U_\nu + (y_{kk} - k y_{ik}) U_k + (y_{kl} - k y_{il}) U_l \tag{7a}$$

$$I_l - \sum_{\substack{\nu=1 \\ \neq k,l}}^{n} y_{l\nu} U_\nu = k I_i + \sum_{\substack{\nu=1 \\ \neq k,l}}^{n} k y_{i\nu} U_\nu + (y_{lk} + k y_{ik}) U_k + (y_{ll} + k y_{il}) U_l. \tag{7b}$$

Diese Gleichungen sind bei Anwendung der Parameterbestimmung nach [1] geeignet zur Berechnung des Steuerparameters k und der Admittanzen $y_{i\nu}$, $y_{j\nu}$ $(\nu = 1, \ldots, n, \neq k, l)$ sowie $y_{k\nu}$ und $y_{l\nu}$ $(\nu \neq k, l, i, j)$. Von den beiden gesteuerten Stromquellen ist also die spannungsgesteuerte Stromquelle vom Standpunkt der Fehlerlokalisierung als günstiges Netzwerkelement anzusehen.

4. Netzwerk mit spannungsgesteuerter Spannungsquelle UUQ, Bild 3

Die Kontrolle des Steuerparameters k ist trivial:

$$k = (U_k - U_l)/(U_i - U_j). \tag{8}$$

Es gilt weiterhin

$$[I_1, \ldots, I_i, \ldots, I_n]^T = (\boldsymbol{Y})_b (\boldsymbol{U}), \tag{9}$$

wobei $(\boldsymbol{Y})_b$ durch Streichungen der k-ten und l-ten Zeile aus $(\boldsymbol{Y})_{NW}$ hervorgeht.

Die Fehlerstrommethode ermöglicht die Unterscheidung von Fehlern in den Zweigen bis auf solche in $y_{k\nu}$ und $y_{l\nu}$ $(\nu = 1, \ldots, n)$: Dafür ist eine Parameterbestimmung erforderlich. y_{kl} kann als nicht vorhanden angesehen werden.

5. Netzwerke mit stromgesteuerter Spannungsquelle IUQ, Bild 4

Es gilt folgendes Gleichungssystem:

$$\begin{bmatrix} I_1 \\ \vdots \\ I_i \\ I_j \\ I_k + I_l \\ \vdots \\ I_n \end{bmatrix} = \begin{bmatrix} & & k & l \\ & i & y & -y \\ (Y)_c + j & & -y & y \end{bmatrix} \begin{bmatrix} U_1 \\ \vdots \\ U_i \\ U_k \\ U_l \\ \vdots \\ U_n \end{bmatrix}, \qquad (10)$$

wobei $(\boldsymbol{Y})_c$ durch Addition der i-ten und der j-ten Spalte ($U_i = U_j$) sowie durch Addition der k-ten und der l-ten Zeile aus $(\boldsymbol{Y})_{NW}$ hervorgeht. Die Fehlerstrommethode läßt die getrennte Fehlererkennung in den Zweigen bis auf y_{kv} und y_{lv} ($v \neq k, l$) zu. Wegen $y_{ij} = y_{ji} = 0$ in einer realen Schaltung ist auch die Wirkung eines Steuerparameterfehlers eindeutig zu erkennen.

6. Netzwerke mit idealem Operationsverstärker

Es wird die Ersatzschaltung (Bild 5) als Nullor [6] gewählt. Das resultierende Netzwerk wird dann durch die Matrix $(\boldsymbol{Y})_c$ beschrieben. Es gelten für die Fehlerortung im Netzwerk die unter Abschn. 5. getroffenen Aussagen.

NaA 8983

Literatur

[1] *Krambeer, H.:* Einige theoretische Grundlagen der Fehleranalyse in analogen Schaltungen. Nachrichtentechnik · Elektronik 27 (1977) 3, S. 117–120

[2] *Krambeer, H.; Westendorff, K.:* Fehlerlokalisierung in analogen Schaltungen auf der Grundlage von Knotenspannungsmessungen. Nachrichtentechnik · Elektronik 32 (1982) 11, S. 460 u. 465–467

[3] *Trick, T. N.; Mayeda, W.; Sakla, A. A.:* Calculations of Parameter from node voltage measurements. IEEE Trans. Circuit Syst. CAS-26 (1979) 7, S. 466–474

[4] *Roytman, L. M.; Swamy, M. N. S.:* One method of the Circuit Diagnosis. Proc. IEEE 69 (1981) 5, S. 661 u. 662

[5] *Gray, P. E.:* Circuit Diagnosis and Thevenin Equivalents. Proc. IEEE 70 (1982) 2, S. 199 u. 200

[6] *Roeder, D.:* Die Knotenleitwertmatrix eines idealen Verstärkers (Nullors). Nachrichtentechn. Z. 24 (1971) 9, S. 465–467

Prof. Dr. sc. techn. *Heinrich Krambeer*, Ingenieurhochschule Wismar, 2400 Wismar, Philipp-Müller-Str.

Zweidraht-Meßdaten-Erfassungssystem mit meßfühlernaher Digitalisierung

M. Seifert, KDT, Magdeburg

Mitteilung aus der Technischen Hochschule Otto von Guericke Magdeburg, Sektion Technische Kybernetik und Elektrotechnik

Deskriptoren: Meßwerterfassung, Meßwertübertragung, digitales System, AD-Wandler, Meßstellenanordnung, Impulszählung

Bei der Meßdatenerfassung für industrielle Anwendungen müssen häufig Meßwerte mehrerer untereinander räumlich weit entfernter Meßfühler zu einer Zentraleinheit bzw. zu Anzeige- oder Registriereinheiten übertragen werden, die sich an verschiedenen Orten befinden können. In vielen Fällen wird hierfür eine Struktur nach Bild 1 eingesetzt. Bei der mehrkanaligen Meßdatenerfassung mit Mikrorechnern sind Systeme nach Bild 2 vorherrschend [1].

Bei großen Entfernungen zwischen den Meßfühlern haben die Strukturen nach den Bildern 1 und 2 den Nachteil eines hohen Material- und Installationsaufwandes für die Verkabelung. Systeme nach Bild 2 eignen sich vorwiegend für kurze Leitungslängen zwischen den Meßfühlern und dem Multiplexer, weil bei längeren Leitungen Fehler infolge Einstreuung elektrischer Störsignale nicht auszuschließen sind.

Durch Digitalisierung der Meßsignale in unmittelbarer Meßfühlernähe und serielle Datenübertragung auf einem einfachen Bus sind beträchtliche Einsparungen an Kabelaufwand und wesentlich größere Flexibilität des Datenaustausches erzielbar. Ein solches System, das mit geringem Hardware- und Softwareaufwand auskommt und leicht erweiterbar ist, wurde an der TH Magdeburg konzipiert und erprobt (Bild 3). Maximal 16 Datensender und maximal 16 Datenempfänger (erweiterbar) sind räumlich an beliebiger Stelle an eine einfache verdrillte Zweidrahtleitung anschließbar. Die Daten eines vorher adressierten Datensenders lassen sich zu einem oder zu mehreren Datenempfängern übertragen. Die Systemsteuerung erfolgt wahlweise durch eine spezielle Ablaufsteuerung oder durch einen Mikrorechner.

Zusätzlich zu den Datensignalen werden über die verdrillte Zweidrahtleitung Adressensignale, die Betriebsspannung (24 V) und bei Bedarf Steuersignale übertragen.

1. Funktion

Die Daten- und die Adressenworte sind impulsanzahlcodiert. Ein Datenwort ist 10 ms lang und besteht bei einer Auflösung

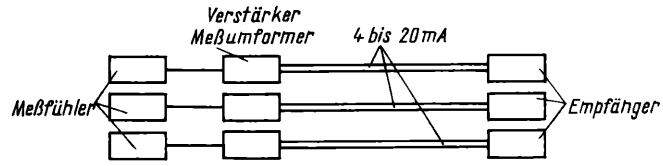


Bild 1. Datenerfassungssystem mit getrennten Kanälen

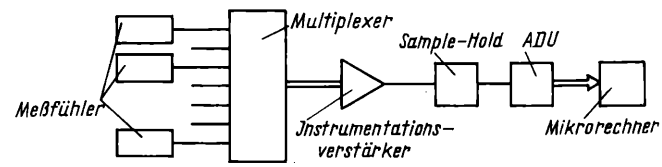


Bild 2. Datenerfassungssystem mit Analogmultiplexer

von 10 bit aus maximal 2^{10} Impulsen von 5 µs Breite und annähernd 24 V Amplitude.

Jeder Datensender und jeder Datenempfänger enthält eine Adressenerkennungsschaltung. Datensender werden mit 50 µs und Datenempfänger mit 100 µs breiten Adressenimpulsen adressiert. Die unterschiedliche Impulsbreite hat den Vorteil, daß der Ausgang eines Datensenders mit den Eingängen *mehrerer* Datenempfänger verbunden werden kann. Falls beispielsweise der Datensender mit der Adresse 2 ein Datenwort zu den Datenempfängern mit den Adressen 1, 3 und 4 senden soll, werden zuerst die drei Datenempfänger und anschließend der Datensender adressiert. Unmittelbar danach wird von der Adressenerkennungsschaltung die Ausgangsstufe des Datensenders aktiviert, wodurch ein Datenwort ausgesendet wird, das zu den vorher adressierten drei Datenempfängern gelangt und dort beispielsweise in einem 10-bit-Dualzähler bis zum Eintreffen eines nächsten Datenwortes zwischengespeichert wird.

Die Adresse des jeweils aktiven Datensenders kann vom Datenempfänger angezeigt werden, wenn er einen Zähler für die Datensender-Adressenimpulse enthält.

Das System benötigt keine getrennten Steuersignale, da der

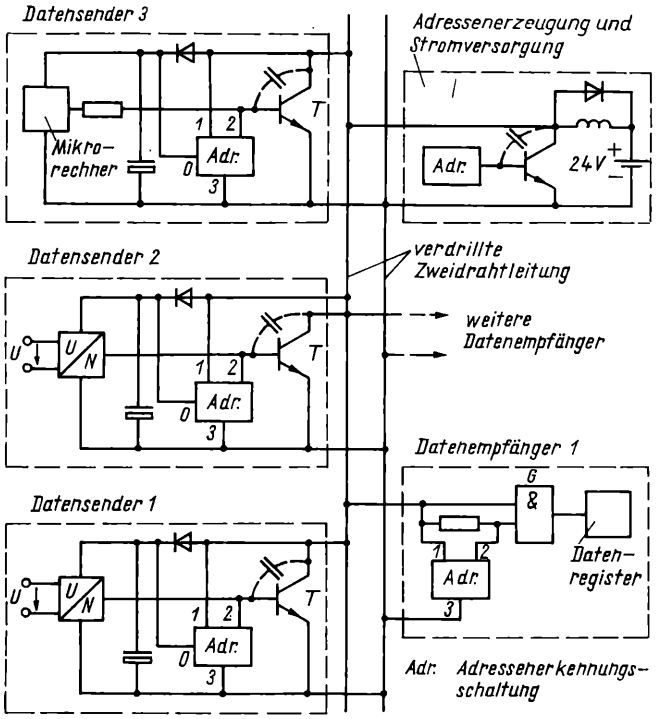


Bild 3. Zweidraht-Meßdaten-Erfassungssystem mit meßfühlernaher Digitalisierung

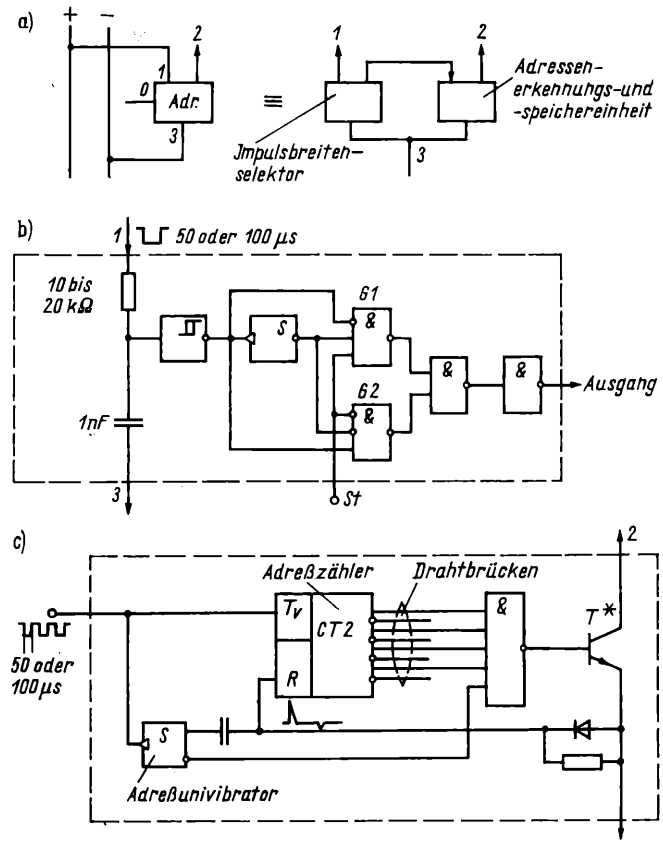


Bild 4. Adressenerkennungsschaltung
a) Prinzip; b) Impulsbreitenselektor;
c) Adressenvergleichs- und -speicherschaltung

adressierte Datensender unmittelbar nach erfolgter Adressierung ein Datenwort aussendet. Es ist auch die zusätzliche Übertragung von Steuersignalen möglich, beispielsweise durch Zuweisung mehrerer Adressen zu einem Datensender bzw. Datenempfänger.

2. Systembeschreibung, Bild 3

Das vom Meßfühler gelieferte Meßsignal wird mit Hilfe eines Spannungs- bzw. Stromfrequenzwandlers [2] [3] in ein impulsanzahlcodiertes digitales Datenwort umgesetzt.

Die Adressenerkennungsschaltung der Datensender ist so dimensioniert, daß sie nur auf Datensender-Adressenimpulse anspricht. Falls die von der Steuereinheit ausgesendete Adresse mit der Adresse eines angeschlossenen Datensenders übereinstimmt, wird dessen Endstufe (offener Kollektor) aktiviert und erzeugt ein Datenwort.

Die Adressenerkennungsschaltung besteht aus einem Impulsbreitenselektor zur Unterscheidung zwischen Datensender- und Datenempfänger-Adressenimpulsen sowie einer Adressenvergleichs- und -speicherschaltung (Bild 4). Im Fall St = H werden Datensender-Adressenimpulse, im Fall St = L werden Datenempfänger-Adressenimpulse zum Ausgang des Impulsbreitenselektors weitergeleitet. Die Adressenvergleichs- und -speichereinheit (Bild 4c) verarbeitet die Ausgangsimpulse des Impulsbreitenselektors. Falls die vom Adressenzähler CT 2 gezählte Anzahl von Adressenimpulsen mit der durch Drahtbrücken einstellbaren Adresse übereinstimmt, wird die Endstufe des Datensenders bzw. der Eingangskreis des Datenempfängers aktiviert. Der wiedertriggerbare Univibrator *S* kippt mit dem ersten Adressenimpuls hin und nach dem letzten Adressenimpuls zurück. Auf diese Weise paßt sich die Adressenerkennungszeit der jeweiligen Länge des Adressenwortes an.

Da auf dem Bus ausschließlich digitale Signale übertragen werden, läßt sich eine galvanisch getrennte Ankopplung der Datensender und Datenempfänger wesentlich problemloser realisieren als bei der analogen Signalübertragung. Daten- und Adressenworte lassen sich mit Optokopplern, die Betriebsspannung mit einem Transverter galvanisch trennen (Bild 5).

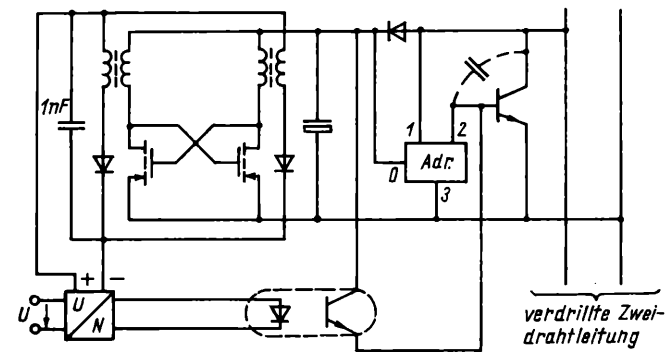


Bild 5. Galvanisch getrennte Signalübertragung

Mehrere Maßnahmen dienen der Vermeidung von Reflexionen und der störsicheren Signalübertragung:

— geringe Flankensteilheit der übertragenen Impulse
— Einschalten von Tiefpaßfiltern vor die Eingänge der Datenempfänger und der Adressenerkennungsschaltungen
— Begrenzung der Reflexionen durch Dioden
— hoher Signalpegel
— verdrillte Signalleitung.

Eingegangen am 15. September 1982 NaA 8984

Literatur

[1] *Seifart, M.:* Analogwerterfassung und -ausgabe mit Mikrorechnern. Nachrichtentech., Elektron. 30 (1980) 11, S. 452–457
[2] *Seifart, M.:* Spannungs- und Stromfrequenzwandler – Überblick und Anwendungen. Nachrichtechn., Elektron. 27 (1977) 6, S. 252–255
[3] *Seifart, M.; Bogk, D.:* Spannungs- und Stromfrequenzwandler mit hoher Linearität und großem Dynamikbereich. Nachrichtentechn., Elektron. 27 (1977) 6, S. 255–258

Prof. Dr.-Ing. *Manfred Seifart*, Technische Hochschule Otto von Guericke Magdeburg, Sektion Technische Kybernetik und Elektrotechnik, 3010 Magdeburg, Postschließfach 124

12-bit-Analog-Digital-Umsetzer mit stufen-paralleler Verarbeitung

V. Brovkov, Dresden

Mitteilung aus der Technischen Universität Dresden, Sektion Informationstechnik

Deskriptoren: Meßelektronik; rechnergestützte Datenerfassung; Analog-Digital-Umsetzung

Die Aufgabe einer schnellen Analog-Digital-Umsetzung wird gegenwärtig hauptsächlich durch die Verwendung von parallelen, parallel-seriell-parallelen oder Stufenumsetzungsverfahren gelöst, wobei eine rein parallele Umsetzung für höhere Umsetzungsauflösungen (10 bis 12 bit) zur Zeit technisch sehr schwer realisierbar ist [1] [2] [3] [4]. Für eine 12-bit-Parallel-Seriell-Parallele-Umsetzung sind u. a. parallele ADUs mit 6 bis 7 bit Auflösung erforderlich [2]. Aus Gründen der leichteren Verfügbarkeit ist die Benutzung eines Stufenverfahrens für eine schnelle und genaue Analog-Digital-Umsetzung als die einfachste Lösung des Problems anzusehen. Die Verwendung diskret oder hybrid realisierter DAUs mit R-2R-Netzwerken sichert eine Umsetzungszeit bei der 12-bit-Auflösung von weniger als 10 µs [3] [4]. Der einfachste Weg, einen schnellen Stufen-ADU unter Verwendung eines industriell gefertigten DAU (z.B. DAC 32, DAC 80) aufzubauen, bereitet einige Schwierigkeiten. Sie liegen darin, daß derartige Schaltkreise einen Ausgangs-OV erfordern, dessen Beruhigungszeiten, bezogen auf den zulässigen Fehler entsprechend einer 12-bit-Auflösung in der Größenordnung von einigen Mikrosekunden liegen. Eine 12-bit-Umsetzung dauert in diesem Fall einige 10 µs. Eine gewisse Beschleunigung der Umsetzung ist durch die Verringerung der Schrittzeit für jeden Folgetakt möglich. Die Beruhigung des OVs während eines Spannungssprunges U besteht aus zwei Zeitphasen (Bild 1) [5] [6]. Am Anfang wird die OVs-Ausgangsspannung U_A mit einer begrenzten Geschwindigkeit ϱ (slow rate) ansteigen, wozu T_s benötigt wird:

$$T_s \approx \frac{U}{\varrho}. \tag{1}$$

Anschließend beginnt die Beruhigungsphase des gegengekoppelten Systems. Für die Ausgangsspannung $U_A(t)$ gilt nach [6]:

$$U_A(t) = U\left[1 - k_1 \exp(-2\pi f_1 t) + k_2 \exp\left(-\frac{t}{\tau}\right)\right] \tag{2}$$

mit

$t > T_s$, $k_1 \approx 1$, $k_2 \ll 1$, $f_1 = f_1$ Frequenz (unity gain bandwith)

woraus unter Vernachlässigung des 3. Terms Gl. (3) folgt:

$$u_{A1} \approx U[1 - \exp(-2\pi f_1 t)]. \tag{3}$$

Der zeitabhängige Restfehler

F_U ($F_U = \Delta U/U$, $\Delta U = u_A(t) - U$)

beträgt:

$$F_U = -\exp(-2\pi f_1 t). \tag{4}$$

Die Beruhigungszeit T_e für einen zulässigen Fehler F_U ist

$$T_e = -\frac{1}{2\pi f_1} \ln F_U. \tag{5}$$

Für den OV K 140 UD 7 mit $\varrho = 10\,\frac{\mathrm{V}}{\mu\mathrm{s}}$, $f_1 = 0{,}7$ MHz bei 12 bit Auflösung ($F_U = 2^{-12}$) und $U_0 = 5$ V Spannungssprung gilt für den ersten Umsetzungsschritt:

$$T_{e0} \approx 8{,}3 \cdot \frac{1}{2\pi f_1} \approx 1{,}9\ \mu\mathrm{s},$$

$$T_{s0} \approx \frac{U_0}{\varrho} = 0{,}5\ \mu\mathrm{s} < T_{e1}.$$

Für den folgenden Umsetzungsschritt des ADUs ist der Spannungssprung um den Faktor 2 kleiner, der absolute Spannungsfehler $U_1 - u_A(t)$ konstant, so daß eine kürzere Beruhigungszeit zu erwarten ist. Für den i-ten Schritt der n-bit-Umsetzung gilt allgemein:

$$T_{ei} = \frac{(n-i)\ln 2}{2\pi f_1} \tag{6}$$

$$T_{si} = \frac{U_0}{\varrho \cdot 2^i} \ll T_{ei}. \tag{7}$$

Die minimale Gesamtumsetzungszeit für 12 bit beträgt:

$$T_{ums} = \sum_{i=0}^{n-1} \frac{(n-i)\ln 2}{2\pi f_1} + nT_L = \sum_{i=0}^{11} \frac{(12-i)\ln 2}{2\pi f_1} + 12\,T_L \approx \frac{8{,}6}{f_1} + 12\,T_L. \tag{8}$$

T_L ist die Verzögerung der Steuerlogik und des Komparators. Eine konventionelle 12-bit-Umsetzung (konstante Taktzeit für alle Stufen) dauert demgegenüber

$$T_{ums} = n \cdot T_{e0} + nT_L \approx \frac{15{,}9}{f_1} + 12\,T_L. \tag{9}$$

Die Verzögerungszeiten der Steuerungslogik und des Komparators sind im Vergleich zu T_e sehr klein. Man erhält aus Gl. (8) eine gegenüber Gl. (9) um etwa den Faktor 2 kleinere Umsetzzeit, muß jedoch einen komplizierteren Taktgenerator vorsehen.

Eine weitere Umsetzungszeitsenkung ist bei einem Umsetzer mit einer Struktur gemäß Bild 2 erreichbar. Dieser Stufen/Parallel-Umsetzer vereinigt in sich die Funktion eines Stufenumsetzers sowie eines Parallelumsetzers. Er enthält einen Parallelumsetzer mit begrenzter Auflösung, eine Steuerungslogik und zwei

Bild 1. Beruhigung eines OVs nach einem Spannungssprung

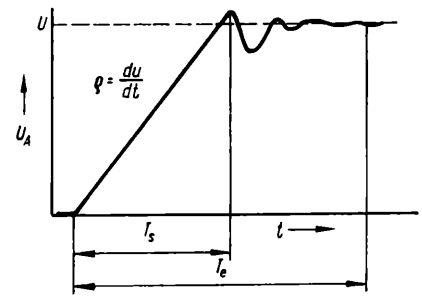
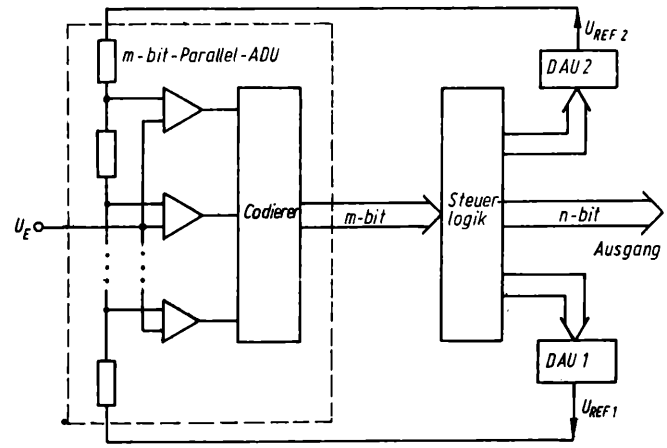


Bild 2. Struktur eines stufen-parallelen ADUs



DAUs, die für den Parallelumsetzer die Referenzspannungen liefern. Besonders hervorgehoben werden soll, daß hier die Referenzspannungen U_{REF1} und U_{REF2} des Unter-ADUs dessen Arbeitsbereich und dessen Empfindlichkeit, somit auch seine Umsetzungskennlinie bestimmen. Am Anfang der Umsetzung sind DAU 1 und DAU 2 so anzusteuern, daß die Referenzspannungen U_{REF1} und U_{REF2} den Grenzen des Umsetzungsbereichs entsprechen (Bild 3a). Das der Eingangsspannung U_E und der durch U_{REF1} und U_{REF2} eingestellten Kennlinie des Unter-ADUs entsprechende Codewort $D1$ gibt die ersten m bits des Ergebnisses. $D1$ zeigt gleichzeitig an, daß die Eingangsgröße zwischen den Spannungen U_{D1} und U_{D1+1} liegt, entsprechend $D1$ und $D1+1$. Werden jetzt DAU 1 entsprechend dem Codewort $D1$ und DAU 2 entsprechend $D1+1$ angesteuert, so stellen die Referenzspannungen $U_{REF1} = U_{D1}$ und $U_{REF1} = U_{D1+1}$ die Kennlinie des Unter-ADUs so ein, daß er im Spannungsbereich U_{D1} bis U_{D1+1} die nächsten m-bit bestimmt (Bild 3b). Gemäß dem nunmehr vorliegenden Ergebnis werden die Referenzspannungen neu eingestellt. Die Schrittzahl einer n-bit-Umsetzung reduziert sich so um den Faktor m gegenüber der üblichen Stufenumsetzung. Für minimale Umsetzungszeit erhält man analog Gl. (8):

$$T_{ums} = \sum_{i=0}^{(n/m)-1} \frac{(n - m\,i)\ln 2}{2\pi f_1} + \frac{n}{m} T_L . \tag{10}$$

Wird $n = 12$ und $m = 3$, dann wird

$$T_{ums} = \sum_{i=0}^{3} \frac{(12 - 3\,i)\ln 2}{2\pi f_1} + 4\,T_L \approx \frac{3{,}3}{f_1} + 4\,T_L . \tag{11}$$

Dieses Prinzip, realisiert mit einem 3-bit-Unter-ADU und 2 DAUs, hat etwa 20 % der Umsetzungszeit einer üblichen Stufenumsetzung. Hierbei ist zu beachten, daß die Fehler des Unter-ADUs sowie DAUs der Auflösung (n-bit) des Umsetzers entsprechen müssen. Weiterhin ist ein spezieller Taktgenerator, bei dem jede Taktzeit entsprechend den Gliedern im 1. Term von Gl. (10) bzw. Gl. (11) zuzüglich T_L festzulegen ist, vorzusehen. In einem Stufen/Parallelen-Umsetzer können die bekannten Überlappungskorrekturverfahren [2] [7] sehr erfolgreich angewandt werden. Dann muß der Unter-ADU während jedes Umsetzungsschrittes eine m-bit-Auflösung mit einem zulässigen Quantisierungsfehler von 0 bis —1 LSB des Unter-ADUs haben. Für die Bildung der Referenzspannung U_{REF2} muß zum vorletzten Bit des Unter-ADU-Ausgangscodes 1 addiert werden. Dabei ist der Auflösungsbereich des Unter-ADUs im folgenden Schritt doppelt breit (Bild 4). Das letzte Bit des Codewortes dient einer digitalen Korrektur des Quantisierungsfehlers durch die dargestellte Überlappungsaddition mit dem ersten Bit des Codewortes des nächsten Umsetzungsschrittes. Wird z.B. eine Eingangsspannung U_E während des ersten Schrittes mit dem Codewort 011 fehlerlos digitalisiert (Bild 4a), hat der Unter-ADU im zweiten Schritt die Kennlinie 1 (Bild 4b), wodurch 000 als Antwort erzeugt wird. Das Ergebnis wird durch eine Überlappungsaddition gewonnen:

$$\begin{array}{r} 011 \\ +\ 000 \\ \hline 01100 \end{array} . \tag{12}$$

Liegt im ersten Umsetzungsschritt ein Quantisierungsfehler von —1 LSB vor, wird ein Codewort 010 erzeugt, dem die Kennlinie 2 des zweiten Umsetzungsschrittes entspricht. Der zweite Schritt liefert 100. Eine Überlappungsaddition führt zum gleichen Ergebnis wie oben:

$$\begin{array}{r} 010 \\ +\ 100 \\ \hline 01100 \end{array} . \tag{13}$$

Das letzte Bit des Ergebnisses kann ebenfalls für eine Überlappungskorrektur im dritten Schritt dienen. Hierbei ist es wesentlich, daß der zulässige Restfehler F_U bei der Einstellung von Referenzspannungen in allen Schritten konstant bleibt und der m-bit-Auflösung des Unter-ADUs entspricht, wodurch alle Schritte die gleiche Zeit benötigen. Für eine n-bit-Umsetzung mit m-bit-Unter-ADU und einer Überlappungskorrektur gilt

$$T_{ums} = \frac{n}{m-1}\left(\frac{m \ln 2}{2\pi f_1} + T_L\right). \tag{14}$$

Bild 3. Umsetzungskennlinien eines stufen-parallelen ADUs

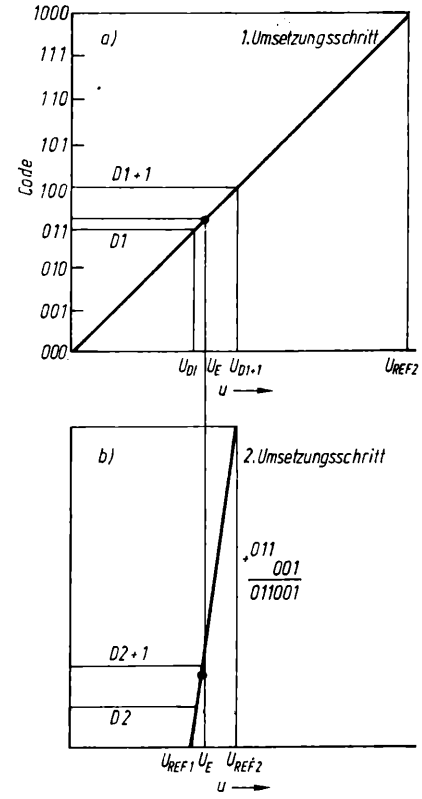
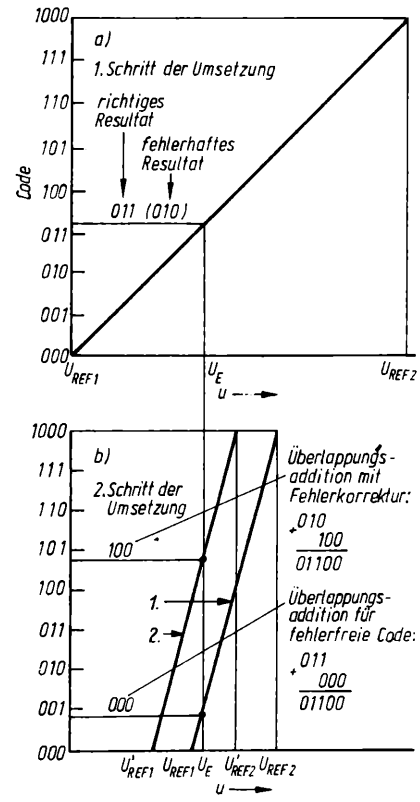


Bild 4. Kennlinien eines stufen-parallelen ADUs mit Überlappungskorrektur



Wird $n = 12$ und $m = 2$, dann wird

$$T_{ums} = 12\left(\frac{2 \ln 2}{2\pi f_1} + T_L\right) \approx \frac{2{,}7}{f_1} + 12\,T_L . \tag{15}$$

Ein Vergleich von Gl. (15) mit Gl. (8) zeigt, daß ein Stufen/Paralleler-Umsetzer mit Überlappungskorrektur eine etwa 6fache Umsetzungszeitsenkung im Vergleich zu üblichen Stufenum-

Fortsetzung auf Seite 14

Algorithmen zur digitalen Effektivwertbestimmung

E.-G. Kranz, KDT, Dresden

Mitteilung aus der Technischen Universität Dresden, Sektion Informationstechnik

Deskriptoren: Meßwertverarbeitung, Effektivwert, Zeitmittelwert, Ensemblemittelwert, Algorithmus, Rechenzeitbedarf

1. Problemstellung

Zur Beschreibung von Zeitfunktionen werden verschiedene Kennwerte und Kennfunktionen benutzt. Dabei hat der Effektivwert für die Behandlung zahlreicher physikalisch-technischer Probleme besondere Bedeutung. Das resultiert vor allem daraus, daß er sich einfach zur Bestimmung energetischer Größen heranziehen läßt. Bei der Mehrzahl der physikalisch realen Signale können die Mittelwerte nach dem Ergodensatz sowohl als statistische Erwartungswerte (Ensemblemittelung) als auch als zeitliche Mittelwerte bestimmt werden [1]. Für den Effektivwert gelten folgende Definitionen:

Zeitliche Mittelung

$$\tilde{x} = \sqrt{\lim_{T \to \infty} \frac{1}{T} \int_0^T x^2(t)\,\mathrm{d}t}\,, \tag{1}$$

Ensemblemittelung

$$\tilde{x} = \sqrt{\int_{-\infty}^{+\infty} x^2\, p(x)\,\mathrm{d}x}\,, \tag{2}$$

für periodische Signale

$$\tilde{x} = \sqrt{\frac{1}{T_p} \int_0^{T_p} x^2(t)\,\mathrm{d}t} \quad (T_p \text{ Periodendauer}). \tag{3}$$

Mit fortschreitender Entwicklung der Elektronik ist auch in der elektronischen Meßtechnik eine wachsende Digitalisierung zu verzeichnen. Insbesondere erschließt der Mikrorechner durch seine Flexibilität, sein geringes Volumen und seinen relativ niedrigen Preis vielfältige Möglichkeiten für die Meßwertverarbeitung. Bild 1 zeigt den prinzipiellen Aufbau einer digitalen Verarbeitungskette für Zeitfunktionen.

Zu diskreten Zeiten im Abstand ΔT werden der Zeitfunktion $x(t)$ Abtastwerte $x(k\Delta T) = x_k$ entnommen. Es besteht nun die Aufgabe, zur Bestimmung des Effektivwertes $\tilde{x}$ die Werte x_k entsprechend den Vorschriften in den Gln. (1) oder (2) zu verarbeiten. Auf geeignete Algorithmen, insbesondere zur Ermittlung des Effektivwertquadrates, wird im folgenden eingegangen.

Fortsetzung von Seite 13

setzern ergibt, obwohl der Aufwand nur etwa zweimal so hoch ist. Damit hat dieses Prinzip besondere Bedeutung für die Realisierung schneller ADUs.

Eingegangen am 15. September 1982 NaA 8985

Literatur

[1] Schnelle Wandler. Technologie und Auswahl. Elektronik Industrie 12 (1981) 11, S. 23–26.

[2] *Pratt, W.:* High linearity and video speed come together in a–d converters. Electronics 53 (1980) 22, S. 167–170

[3] *Schuhmann, R.; Kirsten, T.:* Analog-Digital-Umsetzer für hohe Frequenzen. Wiss. Z. der TU Dresden 30 (1981) 1, S. 87–91

[4] *Dieben, A.; Heymel, G.; Unger, H.:* Schneller Analog-Digital-Wandler für Testzwecke. radio fernsehen elektronik 28 (1979) 10, S. 621–623

[5] *Schilo, W. L.:* Linejnije integralnije shemi. Moskva: Sowjetskoje Radio, 1979

[6] *Veshwant Kamath, B.; Meyer, R. G.; Gray, P. R.:* Relationship Between Frequency Response and Setting Time of Operational Amplifiers. IEEE J. SC-9 (1974), S. 347–352

[7] *Zimmer, M.:* Genaue Analog/Digital-Umsetzung für die Hochfrequenz-Datenerfassung. Elektronik 28 (1979) 6, S. 41–45

Dipl.-Ing. *Vladimir Brovkov*, Technische Universität Dresden, Sektion 9 Informationstechnik, 8027 Dresden, Mommsenstr. 13

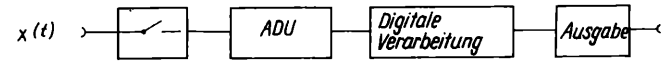


Bild 1. Digitale Verarbeitungskette für Zeitfunktionen

2. Einschränkende Bedingungen

Bei der digitalen Meßwertverarbeitung sind auf Grund technischer Randbedingungen Einschränkungen bzw. Fehler zu erwarten. Dabei sind 2 Hauptgruppen besonders erwähnenswert:

2.1. Fehler auf Grund der endlichen Meßzeit

Da im Rechner nur eine endliche Anzahl N von Meßwerten verarbeitet werden kann, beträgt die Länge des berücksichtigten Zeitfensters

$$T = N\,\Delta T \quad \text{mit} \quad T \leq \frac{1}{2B}. \tag{4}$$

Daraus ergibt sich, daß statt des wahren Mittelwertes nur ein Kurzzeitmittelwert im Sinne eines Schätzwertes für den zu bestimmenden Parameter berechnet wird. Die Güte der Schätzung hängt wesentlich vom Charakter des Signals (z.B. Bandbreite) und der Zeitfensterlänge (Meßzeit) ab. Für die Bestimmung des Effektivwertes von weißem Rauschen mit Bandbegrenzung durch einen idealen Tiefpaß mit der Bandbreite B erhält man z.B. für den mittleren relativen Fehler [2]

$$\frac{D[\tilde{x}]}{\tilde{x}} = \frac{1}{2\sqrt{BT}}. \tag{5}$$

Daraus läßt sich die notwendige Anzahl N der zu verarbeitenden Meßwerte bei vorgegebenem Fehler ermitteln.

2.2. Fehler auf Grund einer endlichen Anzahl von Quantisierungsstufen

Durch die Quantisierung im ADU mit einer Stufenbreite Δx entsteht eine als Quantisierungsrauschen bezeichnete Störkomponente. Ihre Leistung kann für viele praktisch vorkommende Signale mit

$$P_Q = \frac{\Delta x^2}{12} \tag{6}$$

abgeschätzt werden.

Diese Gleichung liefert die Grundlage für eine Abschätzung der notwendigen Anzahl n von Quantisierungsstufen. Außerdem sind die Aussteuerungsgrenzen

$$\hat{x} = n\,\Delta x \tag{7}$$

zu beachten.

3. Algorithmen zur Bestimmung des quadratischen Mittelwertes

3.1. Zeitliche Mittelung

Zur meßtechnischen Bestimmung des quadratischen Mittelwertes wird vorrangig die zeitliche Mittelung benutzt. Unter Beachtung von Gl. (4) ergibt sich als Rechenvorschrift

$$\tilde{x}^2 = \frac{1}{N} \sum_{k=0}^{N-1} x_k^2. \tag{8}$$

Hierbei sind N Multiplikationen zur Bildung der Quadrate auszuführen, was vor allem in Rechnern ohne Parallelarithmetik u.U. erhebliche Rechenzeit beanspruchen kann. Deshalb erscheint es notwendig, auch den Einsatz anderer Algorithmen in Erwägung zu ziehen. Die Division durch N wird zweckmäßigerweise als Verschiebeoperation realisiert.

3.2. Ensemblemittelung

Bei der Definition des Effektivwertes nach Gl. (2) ist die Kenntnis der Verteilungsdichte $p(x)$ Voraussetzung. Sie läßt sich aus

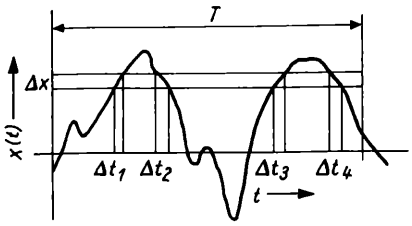


Bild 2. Ermittlung der Verteilungsdichte $p(x)$

der Häufigkeit des Auftretens von Momentanwerten im Bereich

$$x - \frac{\Delta x}{2} \leqq x(t) \leqq x + \frac{\Delta x}{2} \tag{9}$$

ermitteln. Für ergodische Signale kann dazu ein Ensemble von Realisierungen zu einem festen Zeitpunkt oder eine Realisierung, die über eine Zeitdauer T betrachtet wird, herausgezogen werden. In der Meßtechnik wird vorwiegend die zweite Methode angewendet. Gemäß Bild 2 gilt [1]

$$p(x) = \lim_{\substack{T \to \infty \\ \Delta x \to 0}} \frac{1}{\Delta x} \cdot \frac{\Sigma \Delta t_\nu}{T} . \tag{10}$$

Unter Beachtung der in Abschn. 2. genannten Bedingungen liefert die Klassierung von N Meßwerten im Zeitfenster $T = N\Delta T$ als Schätzwert

$$p(x) = \frac{1}{\Delta x} \cdot \frac{N_{xi}}{N} . \tag{11}$$

N_{xi} Anzahl der Abtastwerte mit $x_i - \frac{\Delta x}{2} < x \leqq x_i + \frac{\Delta x}{2}$. Nach [2] sind dabei für $\Delta x \leqq 0{,}2\,\tilde{x}$ die Fehler kleiner als 1 %.
Mit den Gln. (2) und (11) ergibt sich als Rechenvorschrift für den quadratischen Mittelwert

$$\tilde{x}^2 = \sum_{i=-n}^{+n} x_i^2 \frac{N_{xi}}{N} . \tag{12}$$

Falls die Klassierung nach dem Betrag $|x_i|$ vorgenommen wird, kann die Summierung vereinfacht werden:

$$\tilde{x}^2 = \sum_{i=0}^{n} x_i^2 \frac{N_i}{N} . \tag{13}$$

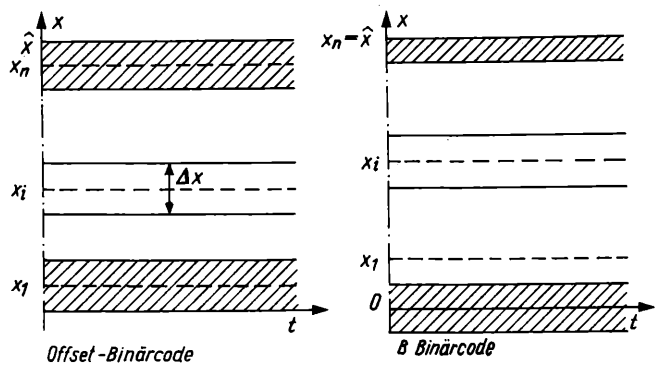
Diese Vorschrift benötigt für $n < N$ weniger Multiplikationen, bedarf dafür aber der Klassierung der N Meßwerte. Die Klassierung kann relativ einfach dadurch realisiert werden, daß der Meßwert (Momentanwert) x_i als Adresse des Klassenzählers i interpretiert wird. Als Codeformen des ADU werden die in Bild 3 dargestellten Fälle betrachtet.
Damit ergeben sich je nach Code folgende Rechenvorschriften:

$$A \quad \frac{x_i}{\hat{x}} = \frac{2i-1}{2n} \qquad \frac{\tilde{x}^2}{\hat{x}^2} = \frac{1}{4n^2 N} \sum_{i=1}^{n} (2i-1)^2 N_i \tag{14}$$

$$B \quad \frac{x_i}{\hat{x}} = \frac{i}{n} \qquad \frac{\tilde{x}^2}{\hat{x}^2} = \frac{1}{n^2 N} \sum_{i=1}^{n} i^2 N_i . \tag{15}$$

Beide Gleichungen enthalten Summen der Form $\Sigma a_i N_i$, deren Koeffizienten a_i mit Hilfe der Gleichung

Bild 3. Codierungsformen der AD-Umsetzung



$$m^2 = \sum_{i=1}^{m} (2i-1) \tag{16}$$

rekursiv berechnet werden können. Die Division kann wieder als Verschiebung realisiert werden. Es werden dann folgende Rechenoperationen benötigt:

Anzahl der Additionen $Z_A = 2n$
Verschiebungen $Z_V = 3n + 2(1 + \mathrm{ld}\, n) + \mathrm{ld}\, N$
Multiplikationen $Z_M = n$.

Eine Umgehung von Multiplikationsoperationen ist möglich, wenn statt der relativen Häufigkeit des Auftretens von $|x_i|$ die relative Überschreitungshäufigkeit N_i^*/N benutzt wird [3]. Sie läßt sich aus der Häufigkeitsverteilung durch Addition ermitteln:

$$\frac{N_i^*}{N} = \sum_{r=i+1}^{n} \frac{N_r}{N} . \tag{17}$$

Damit ergeben sich entsprechend der gewählten Codeform folgende Gleichungen für Fall A bzw. B:

$$\frac{\tilde{x}^2}{\hat{x}^2} = \frac{1}{4n^2}\left[1 + 8 \sum_{i=1}^{n-1} i \frac{N_i^*}{N}\right] \tag{18}$$

$$\frac{\tilde{x}^2}{\hat{x}^2} = \frac{1}{n^2}\left[2 \sum_{i=0}^{n-1} i \frac{N_i^*}{N} + \sum_{i=0}^{n-1} \frac{N_i^*}{N}\right] . \tag{19}$$

Voraussetzung ist dabei die Einhaltung der Aussteuerungsgrenzen. Ersetzt man in den Gln. (19) und (18)

$$\sum_{i=1}^{n-1} i N_i^* = \sum_{i=1}^{n-1} \left(\sum_{k=i}^{n-1} N_k^*\right) \tag{20}$$

erhält man schließlich Rechenvorschriften, die ohne Multiplikationsoperation auskommen. Es ist:

$$\frac{\tilde{x}^2}{\hat{x}^2} = \frac{1}{4n^2}\left[1 + \frac{8}{N} \sum_{i=1}^{n-1} \left(\sum_{k=i}^{n-1} N_k^*\right)\right] \tag{21}$$

mit

$$N_i^* = \sum_{r=i+1}^{n} N_r$$

$$\frac{\tilde{x}^2}{\hat{x}^2} = \frac{1}{n^2 N}\left[2 \sum_{i=1}^{n-1} \sum_{k=i}^{n-1} \sum_{r=k+1}^{n} N_r + \sum_{i=0}^{n-1} \sum_{r=i+1}^{n} N_r\right] . \tag{22}$$

Zur Realisierung kann folgender Bedarf an Rechenoperationen abgeschätzt werden:

Algorithmus nach Gl. (21)
Anzahl der Additionen $Z_A \approx n(n-1)$
Verschiebungen $Z_V = 2\,\mathrm{ld}\, n + \mathrm{ld}\, N + 1$

Algorithmus nach Gl. (22)
Anzahl der Additionen $Z_A = \binom{n+1}{3} + 1$
Verschiebungen $Z_V = 2\,\mathrm{ld}\, n + \mathrm{ld}\, N + 1$.

3.3. Vergleich der Algorithmen

Mit den Rechenvorschriften der Gln. (8), (14), (15), (21) und (22) stehen verschiedene Algorithmen zur Berechnung des quadratischen Mittelwertes zur Verfügung. Je nach Aufgabenstellung ist nun der günstigste Algorithmus auszuwählen. Ein Auswahlkriterium ist die notwendige Rechenzeit. In Bild 4 sind die Ergebnisse einer groben Abschätzung des Rechenzeitbedarfs (normiert auf den Taktzyklus t_c) bei Einsatz der ZVE U 880 dargestellt. Dabei ist angenommen worden, daß eine Multiplikation in den Gln. (14), (15), (21) und (22) 2 ld n Additionen und 2 ld n Verschiebungen entspricht. Für die Zeitmittelung nach Gl. (8) entspricht eine Multiplikation ld n Additionen und ld n Verschiebungen.
Der Rechenzeitbedarf für die Algorithmen, die auf der Ensemblemittelung beruhen, ist für eine Quantisierungsstufenzahl $n > 8$ praktisch unabhängig von der Anzahl N der verwendeten Meßwerte. Hinsichtlich der Rechenzeit ist dabei die Vorschrift nach Gl. (14) oder Gl. (15) am günstigsten. Sie benötigen — wie aus Bild 4 ersichtlich — für Wortbreiten des ADU bis 8 bit bei $N = 1000$ geringere Rechenzeiten als die Zeitmittelung. Für

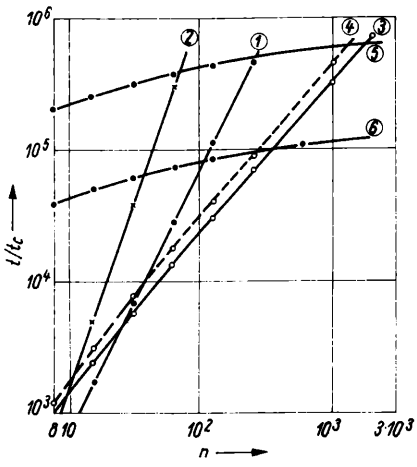


Bild 4. Rechenzeitbedarf (normiert)
1 Algorithmus nach Gl. (21), 2 Algorithmus nach Gl. (22), 3 Algorithmus nach den Gln. (14) und (15), 4 Algorithmus nach den Gln. (14) und (15) bei Koeffizientenberechnung durch Multiplikation, 5 Algorithmus nach Gl. (8) $N = 5000$, 6 Algorithmus nach Gl. (8) $N = 1000$

größere Wortbreiten erscheint jedoch die Zeitmittelung günstiger. Für die Algorithmen nach den Gln. (14), (15) bzw. (21) und (22) sollte neben der Rechenzeit auch die Programmstruktur als Auswahlkriterium beachtet werden. Das wird vor allem bei der Verwendung der Überschreitungshäufigkeit deutlich. Diese Algorithmen verwenden Additionen und erscheinen deshalb auch für Hardware-Realisierungen geeignet. Für große Werte von n sind sie aber nicht besonders zeiteffektiv. Aus Bild 4 ist ebenfalls erkennbar, daß bei Verwendung der Überschreitungshäufigkeit die AD-Wandlung nach dem Offset-Binärcode vorteilhaft ist.

4. Algorithmen zum Radizieren

Sowohl für die Bestimmung des Effektivwertes als Zeitmittelwert als auch als Ensemblemittelwert ist die Operation des Radizierens notwendig. Hierfür werden in Rechenanlagen iterative Algorithmen verwendet. Neben solchen, die Divisionsoperationen benötigen, sind für die Meßwertverarbeitung mit Mikrorechnern besonders die von Interesse, die auf wiederholter Subtraktion mit Restrückstellung beruhen [4] oder sich der Gleichung

$$a^2 = \sum_{k=1}^{a} (2k - 1) \tag{23}$$

bedienen [3] [4]. Auch hier ist zu verzeichnen, daß ein Kompromiß zwischen Zeiteffektivität und Einfachheit der Struktur geschlossen werden muß.

5. Schlußbemerkungen

Die Bestimmung des Effektivwertes von Zeitfunktionen kann sowohl als Zeitmittelwert als auch als statistischer Erwartungswert erfolgen. Dabei sind hinsichtlich der Meßzeit und der Amplitudenauflösung signalspezifische Randbedingungen einzuhalten. Zur digitalen Meßwertverarbeitung stehen mehrere Algorithmen zur Verfügung, deren Auswahl problemorientiert erfolgen muß. Ein wesentliches Auswahlkriterium ist die benötigte Rechenzeit. Für die Realisierung erscheinen aber auch Strukturfragen bedeutungsvoll.

Eingegangen am 15. September 1982 NaA 8986

Literatur

[1] *Kraak, W.:* Signalanalyse. Studienmaterial für die Weiterbildung W 9–153. Techn. Universität Dresden 1977
[2] *Bendat, J. S.; Piersol, A. G.:* Random data: Analysis and measurement procedures. New York: John Wiley 1971
[3] *Deist, F.; Kitai, R.:* Digital transfer voltmeters: principles and error characteristics. Proc. IEE 110 (1963) 10, S. 1887–1904
[4] *Schwarz, W.; Meyer, G.; Eckhardt, D.:* Mikrorechner. Berlin: VEB Verlag Technik 1980

Dr.-Ing. *Ernst-Georg Kranz*, Technische Universität Dresden, Sektion 9 Informationstechnik, 8027 Dresden, Mommsenstr. 13

Digitale Funktionserzeugung nach dem Tabellenverfahren

G. Uhlig, KDT, Dresden

Mitteilung aus der Technischen Universität Dresden, Sektion Informationstechnik

Deskriptoren: Elektronische Meß- und Prüftechnik, Signalgenerierung, digital; Mikrorechneranwendung

Die Erzeugung von programmierbaren Signalfunktionen ist für die automatische Meß- und Prüftechnik eine grundlegende Voraussetzung. Für die Generierung von in Frequenz, Amplitude, Signalform und anderen Parametern programmierbaren Funktionen bieten sich digitale Syntheseverfahren an. Diese Prinzipien sind durch die schnelle Entwicklung der Schaltkreistechnik ökonomisch realisierbar geworden.
Das Tabellenverfahren hat gegenüber den rechnenden Verfahren, den Zählverfahren und dem Summationsverfahren die größte Bedeutung erlangt [1].

1. Grundprinzip des Tabellenverfahrens

Charakteristisch für das Tabellenverfahren ist, daß die Stützwerte der zu generierenden Funktion in einem Signalwiederholspeicher ständig vorhanden sind. Der Speicher kann sowohl ein RAM als auch ein ROM sein. Von einer Baugruppe, die die notwendigen Adressen bereitstellt, gesteuert, werden die Stützwerte der Funktion zyklisch ausgegeben. Der sich anschließende Digital-Analog-Wandler setzt die binär codierten Daten in analoge Spannungen um. Dieses Signal kann, falls erforderlich, ein Glättungsfilter durchlaufen und entsprechend den Anwendererfordernissen verstärkt werden (Bild 1).
Für die Variation der zeitlichen Folge der Bereitstellung der

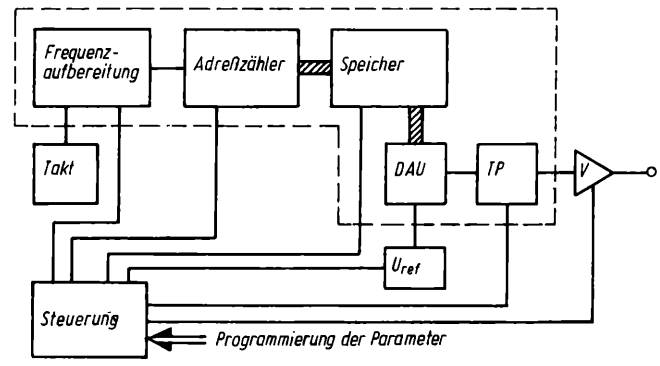


Bild 1. Blockschaltbild eines programmierbaren digitalen Funktionsgenerators nach dem Tabellenverfahren

Stützwerte einer programmierbaren Funktion bestehen zwei grundsätzliche Möglichkeiten, die aus Gl. (1) abzulesen sind:

$$x(t) = F\left(\frac{2\pi}{\Delta\varphi} \cdot \Delta T\right) \tag{1}$$

$\Delta\varphi$ Phasenwinkel zwischen zwei Stützwerten; ΔT Zeitabstand zweier Stützwerte.
Der Begriff Phasenwinkel, der eigentlich nur für sinusförmige Funktionen erklärt ist, soll hier als Abstand zweier Stützwerte in der Folge des Signalformspeichers verstanden werden.

— Der Phasenunterschied zwischen zwei aufeinanderfolgenden Stützwerten ist konstant, die Frequenzänderung geschieht durch Variation der Auslesefrequenz (Bild 2).
Ein Vorteil des Prinzips ist, daß, unabhängig von der Taktfrequenz, stets die gleiche Anzahl von Stützwerten benutzt wird. Das Signalspektrum hat somit, relativ zur Taktfrequenz, immer die gleichen Eigenschaften, ein Umstand, der bei der Erzeugung nichtsinusförmiger Funktionen wesentlich ist. Darüber hinaus kann die Adreßbereitstellung unkompliziert realisiert werden.
Der entscheidende Nachteil des Verfahrens besteht darin, daß besonders dann, wenn nur wenige Stützwerte je Periode ausgegeben werden, ein Mitlauffilter benutzt werden muß. Weiterhin ist die variable Taktfrequenz zur Ansteuerung der Schaltung in einer vorgeschalteten Baugruppe zu erzeugen.

— Der Zeitabstand zwischen zwei aufeinanderfolgenden Stützwerten ist konstant. Durch Vergrößerung bzw. Verkleinerung des Phasenunterschieds kann die Ausgangsfrequenz beeinflußt werden (Bild 3).

Der wesentliche Vorteil ist, daß ein Tiefpaßfilter mit fester Eckfrequenz benutzt werden kann, da das Störspektrum, das sich links und rechts von Vielfachen der Taktfrequenz befindet, ebenfalls konstant ist. Adreßerzeugung und Taktfrequenzaufbereitung können günstig kombiniert werden.
Als Nachteil muß genannt werden, daß der Hardwareaufwand für die programmierbare Adreßbereitstellung relativ groß ist. Falls der eingestellte Phasenschritt in keinem ganzzahligen Verhältnis zur Anzahl der Abtastwerte je Periode steht, können Signalverfälschungen auftreten [2].

Bild 2. Frequenzvariation bei phasenkonstanter Abtastung

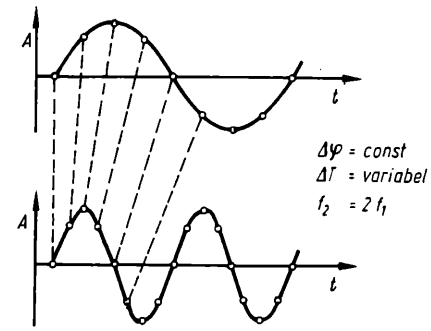


Bild 3. Frequenzvariation bei zeitkonstanter Abtastung

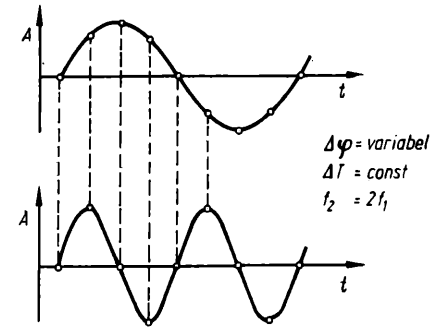
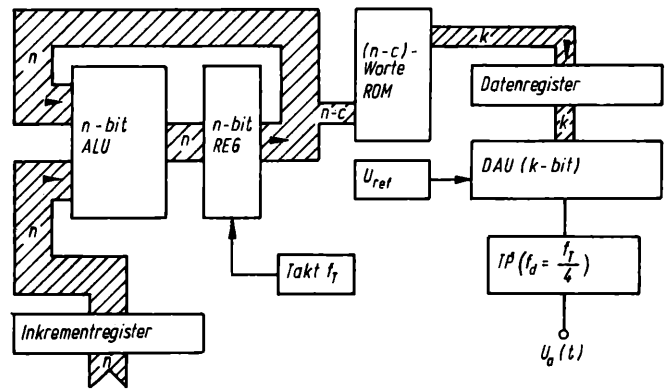


Bild 4. Blockschaltbild eines Phasenakkumulationsgenerators ▼



2. Realisierungsmöglichkeiten

2.1. Adreßbereitstellung durch Zähler

Dieses Verfahren ist ein Vertreter des phasenkonstanten Generierungsprinzips. Die wichtigsten Vor- und Nachteile wurden im vorangegangenen Abschnitt bereits genannt. Die Nutzung eines Binärzählers zur Bereitstellung der Stützwertadresse sichert, daß zyklisch alle vorhandenen Stützwerte gelesen werden. Die maximale Ausgangsfrequenz f_a wird von den Hardwareeigenschaften (maximale Taktfrequenz des Zählers, Speicherzugriffszeit, Wandlungszeit des DAU) und der Anzahl der Stützwerte je Periode (N) bestimmt:

$$f_a = f_T \cdot \frac{1}{N}. \tag{2}$$

Die Verzögerungszeiten der Baugruppen Zähler, Speicher und DAU müssen kleiner als die Taktperiode sein ($T_T = 1/f_T$). Aus Gl. (2) ist zu ersehen, daß man einen Kompromiß zwischen der Stützwertanzahl und damit der Signalqualität und der oberen Frequenzgrenze eingehen muß. Bei nichtsinusförmigen Funktionen sollte die Anzahl der Stützwerte 100 nicht wesentlich unterschreiten, da sich sonst die Signaleigenschaften verschlechtern (Linearität von Ausstiegen, Flanken).
Die Erzeugung des programmierbaren Ansteuertaktes kann mit analogen, hybriden (PLL) oder digitalen Verfahren, wie sie im folgenden beschrieben werden, erfolgen.

2.2. Phasenakkumulationsverfahren

Das Phasenakkumulationsverfahren vereinigt in sich die Baugruppen Frequenzaufbereitung und Adreßzähler. Die Hauptkomponenten sind im Blockschaltbild (Bild 4) dargestellt.
In einem Akkumulator, der in der Lage ist, eine modulo-2-Addition auszuführen, werden eine programmierbare Konstante, die man als Inkrement bezeichnet, und das Ergebnis der vorangegangenen Addition, das in einem Register zwischengespeichert war, addiert. Das geschieht in einem konstanten Zeitabstand ΔT. Die Ausgangsbitbelegung des Akkumulators bzw. des Registers stellt die Adreßinformation dar. Falls nur eine programmierbare Impulsfrequenz benötigt wird, ist der Überlaufimpuls des Akkumulators als Ausgangssignal verwendbar [3].
Im Ergebnis der zyklischen Addition des Inkrements i stehen am Ausgang ständig neue Stützwerte der zu generierenden Funktion bereit. Bei einem Gesamtumfang M des Akkumulators erhält man

$$f_a = \frac{i}{M} \cdot f_T. \tag{3}$$

Da $g = M/i$ im allgemeinen keine ganze Zahl ist, wird der Akkumulator vor jedem Überlauf unterschiedlich weit gefüllt sein. Die Folge ist, daß die Übertragsimpulse einen ungleichmäßigen Abstand haben.
Von einem n bit breiten Akkumulator werden nur die höherwertigen n-c bits zur Adressierung des Speichers benutzt. Diese Einschränkung ist sowohl wegen der Realisierbarkeit des Speichers erforderlich als auch zweckmäßig. Aus der Störabstandsbetrachtung entsprechend Gl. (4) für die Sinusfunktion ergibt sich, daß die Adreßbreite des Wiederholspeichers nur 2 bit größer als die Umsetzbreite des DA-Wandlers sein muß:

$$\frac{S}{N} = 20 \lg \left(\sqrt{\frac{\pi^2}{3} \cdot 2^{-2(n-c)} + \frac{2}{3} \cdot 2^{-2k}} \right)^{-1}. \tag{4}$$

Gleichung (4) gilt, wenn ein Glättungsfilter benutzt wird, das alle Signalanteile größer als $f_T/2$ sperrt. Das Inkrement kann im Bereich von 1 bis etwa $M/4$ gewählt werden. Die obere Grenze ergibt sich aus den praktisch realisierbaren Filtereigenschaften. Auf Grund der Aufteilung der Akkumulatorausgänge besteht dann, wenn Bits im höher- und im niederwertigen Teil des Akkumulators gesetzt sind, die Möglichkeit, daß der Störabstand bis zu 6 dB gegenüber dem Ergebnis von Gl. (4) abnimmt.
Ein Phasenakkumulationsgenerator kann sowohl hardwaremäßig als auch mit einem Rechenprogramm realisiert werden. Die Softwarelösung ist nur für Frequenzen bis etwa 10 kHz geeignet. Sollen höhere Frequenzen generiert werden, so sind CMOS, TTL oder ECL-Varianten vorzuziehen.
Die Verarbeitungsbreite liegt meist zwischen 20 und 40 bit, je nach gewünschter Auflösung. Für den Akkumulator sind am

besten Adder (z.B. 4-bit-Volladder 7483 ≙ K 155 IM 3) und Recheneinheiten (ALU Arithmetic-Logic Unit, z.B. 4-bit-ALU 74181 ≙ K 155 IP 3) geeignet. Für besonders hohe Taktfrequenzen stehen diese Typen auch in Schottky-TTL und ECL zur Verfügung.

Bei Verwendung von ALU-Schaltkreisen besteht die Möglichkeit, mit einem „schnellen Übertrag" zu arbeiten und so die Rechenzeit zu verringern. Die maximal verarbeitbaren Taktfrequenzen liegen bei 10 MHz (TTL), 20 MHz (STTL) und 40 MHz (ECL). Der Stützwertspeicher muß in der Lage sein, die Bedingung $f_T = 1/t_{zu}$ zu erfüllen. Die Zugriffszeiten t_{zu} für in der DDR verfügbare Speichertypen betragen für den EPROM U 555 sowie für den RAM U 202 etwa 450 ns und verringern sich beim RAM U 215 auf unter 100 ns. Für noch schnellere Systeme wären Speicher auf der Basis von TTL- bzw. ECL-Schaltkreisen erforderlich.

Die gleichen Zeitbedingungen gelten für den DA-Wandler. Auch hier sind Umsetzzeiten von weniger als 100 ns nur mit vergrößertem Aufwand erreichbar.

Aus den Betrachtungen zur Realisierung des Phasenakkumulationsverfahrens kann man ableiten, daß dieses Prinzip gegenwärtig für Taktfrequenzen bis etwa 10 MHz einsetzbar ist. Die Auflösung hängt nur von der Arbeitsbreite des Akkumulators ab und ist prinzipiell nicht begrenzt. Das Verfahren ermöglicht die Generierung von Sinussignalen im Frequenzbereich von 1 μHz und weniger bis etwa 2,5 MHz und einem Klirrfaktor von 0,1 % bis 1 %.

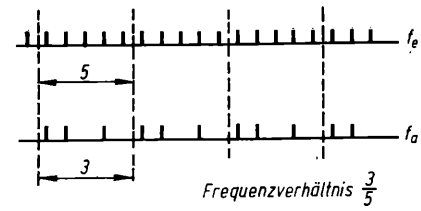
2.3. Binary-Rate-Multiplier-Prinzip

Dieses Verfahren ist ebenfalls durch eine rein digitale Funktionsweise charakterisiert.

Die Programmierbarkeit des Teilverhältnisses wird dadurch realisiert, daß bei einer kontinuierlichen Eingangsimpulsfolge Impulse aus- bzw. eingeblendet werden. Um zu erreichen, daß die entstandenen spektralen Komponenten einen möglichst kleinen Abstand voneinander haben, wählt man die Pause zwischen den Impulsen zu n und $n+1$ Impulsabständen (vgl. Bild 5, $n = 0$). Die Störunterdrückung ist durch eine nachfolgende ganzzahlige Teilung zu verbessern [4].

Der Aufbau von Teilern nach dem Rate-Multiplier-Prinzip ist mit Schieberegistern, Flipflops, Zählern oder Speicher möglich. Als TTL-Schaltkreis steht mit dem Typ 7497 ≙ K 155 IE 8 ein synchroner Teiler zur Verfügung, dessen Teilverhältnisse im Bereich von 64 : 1 bis 64 : 63 programmierbar sind.

Bild 5. Prinzip des Rate-Multiplier-Verfahrens



Dieses digitale Verfahren ist besonders für die Erzeugung variabler Taktfrequenzen geeignet, wie sie z.B. zur Ansteuerung eines Adreßzählers nach Abschn. 2.1. notwendig sind. Bei Realisierung in TTL-Technik können Eingangsfrequenzen bis 30 MHz verarbeitet werden.

3. Schlußfolgerungen

Die genannten digitalen Frequenz- bzw. Funktionsgenerierungsprinzipien wurden untersucht und miteinander verglichen. Sie sind geeignet, in einem weiteren Frequenzbereich, der nach unten durch die Verarbeitungsbreite und nach oben durch die Verzögerungszeiten von Adreßbereitstellung, Signalformspeicher und DAU bestimmt ist, zur Signalerzeugung eingesetzt zu werden. Besonders im Zusammenwirken mit einem Mikrorechner, der aus der gewünschten Frequenz die Programmierdaten für die Synthesebaugruppe errechnet, sind effektive Lösungen möglich. Darüber hinaus ermöglicht die Rechnersteuerung den Schritt zum „intelligenten Meßgerät", d.h. zu einem Gerät, das Grundprozesse der automatischen Meß- und Prüftechnik selbst steuert. Mit einem Interface-Anschluß ist die Zusammenarbeit mit anderen Komponenten in einem System möglich.

Eingegangen am 15. September 1982 NaA 8987

Literatur

[1] *Frühauf, U.; Geiger, L.; Kranz, E.-G.; Kunze, G.:* Zur integrierten Signalgenerierung und -verarbeitung in Meßgeräten. Nachrichtentech., Elektron. 31 (1981) 1, S. 19 – 21

[2] *Hosking, R. H.:* Direct Digital Frequency Synthesis. IEEE Intercon. Tech. Papers 1973 34/1, S. 1 – 6

[3] *Tierney, J.; Rader, Ch.; Gold, B.:* A Digital Frequency Synthesizer. IEEE Trans. AU-19 (1971) 1, S. 48 – 56

[4] *Becker, G.:* Grundlagen der aperiodischen und quasiperiodischen Frequenzsynthese. Frequenz 27 (1973) 9, S. 249; 10, S. 279

Dipl.-Ing. *Günther Uhlig*, Technische Universität Dresden, Sektion 9 Informationstechnik, 8027 Dresden, Mommsenstr. 13

Geräte zur phonetischen Eingabe und Ausgabe – gegenwärtiger Stand und Entwicklungstendenzen

K. Püschel, Dresden

Mitteilung aus dem VEB Robotron, Zentrum für Forschung und Technik, Dresden

Deskriptoren: Spracherkennung, Sprachsynthese

Anfang der 70er Jahre vollzog sich international ein sprunghafter Anstieg von Forschungs- und Entwicklungsarbeiten auf dem Gebiet der phonetischen Eingabe und Ausgabe (automatische Spracherkennung und Sprachsynthese). Ein wesentlicher Anstoß hierfür resultierte aus der Tatsache, daß sich eine Kluft zwischen der immer höheren Leistungsfähigkeit und Anwendungsbreite von Computersystemen einerseits und den verfügbaren Möglichkeiten einer effektiven Mensch-Maschine-Kommunikation andererseits abzeichnete. Seit Ende der 70er Jahre werden von mehreren führenden Elektronik-Konzernen sowie einigen speziell auf Sprachkommunikationsgeräte orientierten Ingenieurfirmen der USA, Japans und der BRD Geräte für die phonetische Eingabe und Ausgabe in einer breiten Palette unterschiedlicher Leistungsparameter angeboten. Durch den Einsatz spezieller LSI-Schaltkreise wurden für die Sprachausgabe bereits billige Massenanwendungen als Konsumgüter (Uhren, Taschenrechner, Lernhilfen, elektronische Wörterbücher oder Spielzeuge mit Sprachausgabe) auf den Markt gebracht; bei der Spracheingabe steht dieser Schritt unmittelbar bevor. Die breite Anwendung derartiger Geräte auf technischem Gebiet (Mensch-Maschine-Kommunikation im weitesten Sinne) steht gegenwärtig international noch in der Anfangsphase; jedoch prognostiziert ein USA-Marktforschungsinstitut allein für Spracheingabegeräte schon einen Umsatz von etwa 150 Mio. Dollar im Jahre 1983 und einen Anstieg auf etwa 1 Mrd. Dollar für 1988 [1].

Erste Anwendungserfahrungen haben gezeigt, daß durch die Sprachkommunikation keineswegs die herkömmlichen Eingabe- und Ausgabegeräte der Computertechnik ersetzt werden können; andererseits lassen sich aber auch neuartige Anwendungs-

möglichkeiten erschließen, die aus einigen Besonderheiten der Sprache resultieren:

— Sprachinformationen können vom Menschen aufgenommen bzw. abgegeben werden, während Hände und Gesichtssinn gleichzeitig mit anderen Arbeitsaufgaben beschäftigt sind
— auf Grund der räumlichen Schallausbreitung bzw. durch Nutzung eines Telefon- oder Sprechfunknetzes ergibt sich eine hohe Mobilität des Nutzers
— die Sprachkommunikation ist unabhängig von den Beleuchtungsverhältnissen möglich.

Demgegenüber stehen solche Nachteile, wie die Störbeeinflussung durch Umgebungsgeräusche sowie die Tatsache, daß die Sprache ein flüchtiger Vorgang ist, so daß für eine Archivierung oder Wiederholung der Originaldaten ggf. besondere Maßnahmen vorzusehen sind.

Anliegen des vorliegenden Beitrags ist es, einen Überblick zu international angebotenen Lösungen zu geben sowie Entwicklungstendenzen aufzuzeigen. In einem abschließenden Kapitel wird das vom VEB Robotron ZFT in Forschungskooperation mit der TU Dresden, Sektion Informationstechnik, [5] entwickelte Sprach-Eingabe- und Ausgabegerät SEA 1600, vorgestellt.

1. Überblick zur phonetischen Eingabe

Die prinzipielle Struktur einer Einrichtung zur Spracherkennung ist in Bild 1 dargestellt. Für die Merkmalextraktion (Analyse), d.h. die Ableitung von geeigneten Parametern aus der Schalldruck-Zeitfunktion, die im nachfolgenden Entscheidungssystem eine hinreichend signifikante Unterscheidung zwischen den zu erkennenden sprachlichen Einheiten ermöglichen, ist heute eine Vielzahl unterschiedlicher Verfahren bekannt, z.B. die Parallelfilteranalyse (realisiert mit analogen oder digitalen Bandpässen), die Fourier-Analyse (FFT-Prozessor), die LPC-analyse (Linear Predictive Coding) oder die Nulldurchgangsanalyse. Kombinationen und Ableitungen aus diesen Merkmalen wurden ebenso experimentell untersucht wie zusätzliche eindimensionale Merkmale des Sprachsignals (z.B. Energie- und Grundfrequenzverläufe). In den gegenwärtig auf dem Markt angebotenen Spracherkennern wird vorrangig eine Parallelfilter-Analyse (vielfach realisiert durch spezielle LSI-Schaltkreise) angewendet. Die in jüngster Zeit erfolgte Entwicklung programmierbarer digitaler Signalprozessor-Chips mit Festkomma-Multiplikationszeiten von etwa 250 ns für 8- oder 16-bit-Operanden bietet völlig neue Möglichkeiten, eine komplette digitale Filterbank [6] bzw. einen FFT- oder LPC-Analysator mit wenigen Schaltkreisen zu realisieren. Generell ist der Trend abzusehen, daß für die Merkmalgewinnung in zukünftigen Geräteentwicklungen einige wenige Spezialschaltkreise verfügbar sein werden, die gleichzeitig auch in anderen Geräten, z.B. der Nachrichten- oder Automatisierungstechnik, in großen Stückzahlen angewendet werden können.

Die Aufgabe des Entscheidungssystems besteht darin, unter Anwendung eines geeigneten Ähnlichkeitskriteriums von allen möglichen Referenzvektorfolgen diejenige zu ermitteln, die mit der einlaufenden Merkmalvektorfolge die größte Übereinstimmung hat. Die dabei angewendeten Verfahren sind sehr vielfältig, ihr Einsatz hängt in starkem Maße vom Umfang des Vokabulars sowie der Art der zu erkennenden sprachlichen Einheiten (Einzelworterkennung oder Erkennung fließend gesprochener Sätze mit einer geeignet definierten Sprachstruktur) ab.

Die Ähnlichkeits- oder Abstandskriterien bestimmen in entscheidendem Maße den für die gesamte Erkennung erforderlichen CPU-Aufwand, so daß man insbesondere bei kleinen Erkennern bestrebt ist, möglichst einfache Algorithmen zu benutzen, z.B. einen Bitmustervergleich bei rein binären Vektoren oder die City-Block-Distanz [5], die keine multiplikativen Verknüpfungen der Vektorkomponenten erfordert. Um insbesondere bei größeren Vokabularen noch ausreichende Erkennungsraten zu erzielen, sind jedoch aufwendigere Algorithmen, die vielfach in einem engen Zusammenhang zu der Art der benutzten Merkmale stehen müssen, anzuwenden.

Für die Organisation des Vergleichs lassen sich zwei im Aufwand und der Leistungsfähigkeit unterschiedliche Verfahren angeben:

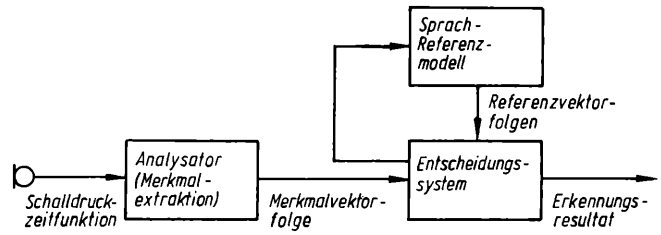


Bild 1. Prinzip der automatischen Spracherkennung

— Kompletter Vergleich der auf eine konstante Vektoranzahl komprimierten Eingangsvektorfolgen mit fest gespeicherten Referenzmustern aller zu erkennenden Wörter (anwendbar bei Einzelworterkennung mit Vokabularen bis etwa 50 Wörter)
— Vergleich auf der Ebene einzelner Merkmalvektoren verbunden mit der Möglichkeit; die durch Ausspracheunterschiede entstandenen nichtlinearen Abweichungen in der Zeitachse zwischen einlaufender Realisierung und den Referenzmodellen durch Anwendung der dynamischen Optimierung zu beseitigen (anwendbar bei Einzelworterkennung mit großem Vokabular sowie der Erkennung fließender Sprache). Ausgehend von einem Strukturmodell der zu erkennenden Wörter bzw. kompletter Sätze sowie einer gespeicherten Menge von Referenzvektoren werden die zulässigen Referenzvektorfolgen parallel zum Voranschreiten des Vergleichsprozesses und ggf. abhängig von bereits getroffenen Entscheidungen generiert (Referenzfolgeverfahren [5]). Bei der Erkennung fließender Sprache, insbesondere dann, wenn von der natürlichen Sprache ähnliche Konstruktionen erkannt werden sollen, kann dieses Strukturmodell mehrere Generierungsebenen, ausgehend von einer linguistischen Ebene bis hin zu einer phonemähnlichen Ebene, umfassen.

Die erforderlichen Referenzinformationen werden durch einen für ein bestimmtes Vokabular einmalig durchzuführenden Lernvorgang ermittelt bzw. bei linguistischen Informationen in Form geeignet formalisierter Regeln vorgegeben. Die gegenwärtig bekannten Lösungen für Spracherkenner sind (abgesehen von Geräten für sehr kleine Vokabulare) noch stark sprecherabhängig, so daß zum Erzielen optimaler Erkennungsraten ein erneutes Anlernen auf jeden unbekannten Sprecher erfolgen muß. Diese sprecherspezifischen Referenzinformationen können gespeichert und dadurch jederzeit für eine Nutzung des Erkenners von dem betreffenden Sprecher aktiviert werden.

Gegenwärtig befinden sich auf dem Markt bereits mehr als 10 Spracherkennertypen von Firmen der USA, Japans und der BRD für Vokabulare von etwa 20 bis zu einigen 100 Wörtern und mit Erkennungsraten in der Größenordnung von 98 bis 99 %. Dabei dominieren Geräte für die Einzelworterkennung, d.h. es sind deutliche Wortpausen bei der fortlaufenden Eingabe mehrerer

Tafel 1. Zusammenstellung der wichtigsten auf dem Markt befindlichen Spracherkenner (nach Literaturangaben und Firmenschriften)

Typ	Wortschatz	Erkennungsrate	Preis	Bemerkungen
VERBEX Mod. 1800	≈ 200	99 %	65 T $	bis zu 5 Wörter fließend eingebbar
DP-100	≈ 200	99 %	65 T $	
T-500	bis 220	99 %	12 T $	Gerät 495 mm × 133 mm × 660 mm, 23 kg
CSE 1060	40 bis 250	99 %	37,4 bis 51,5 TDM	Gerät 450 mm × 150 mm × 660 mm, 25 kg
AURICLE 1	40	99 %	2,5 T $	1 Leiterplatte, mehrere Spezial-LSI-Schaltkreise
MIKE 4725	32	97 %	3,5 T $	1 Leiterplatte, mehrere Spezial-LSI-Schaltkreise
VCR 100-1	bis 100	99 %	100 $	2-Chip-Lösung, sprecherspezifische Referenzdaten auf ROM

Wörter erforderlich; die fließende Eingabe kurzer Wortfolgen bis zu etwa 5 Wörtern ist bei einigen, jedoch sehr teuren Geräten möglich. Tafel 1 gibt einen aus der Literatur und aus Firmenschriften zusammengestellten Überblick zu den wichtigsten in Produktion befindlichen Geräten. Spracherkenner für fließend gesprochene Sätze mit einer der natürlichen Sprache angenäherten Satzbildung sowie für Vokabulare bis zu einigen 1000 Wörtern befinden sich gegenwärtig international noch im Stadium der Grundlagenforschung. Die derzeit experimentell erreichten Erkennungsraten in der Größenordnung von 90% sowie der notwendige Aufwand (mehrfache Echtzeit, bei Einsatz von Großrechnern) erfordern noch umfangreiche Forschungsarbeiten, insbesondere zu folgenden Problemen:

— Entwicklung geeigneter linguistischer Sprachmodelle, um eine effektive Erkennungssteuerung zu realisieren, die für die sichere Erkennung großer Vokabulare in fließender Sprache unumgänglich ist.

— Ableitung leistungsfähiger Merkmale höherer Ordnung, die durch geeignete Algorithmen aus den bekannten primären Merkmalen gewonnen werden und einerseits eine gute Unterscheidbarkeit der Sprachbausteine, andererseits eine möglichst geringe Sprecherabhängigkeit haben sollen. Hierbei werden auch Arbeiten zur Psychoakustik und Hörphysiologie neue Lösungen aufzeigen können.

— Entwicklung von Lern- und Adaptationsverfahren, die das Einstellen eines Spracherkenners auf ein neues Vokabular sowie dessen Anpassung an die Besonderheiten eines unbekannten Sprechers (z.B. lediglich durch Sprechen einiger Codeworte) mit einem praktikablen Zeitaufwand ermöglichen.

2. Überblick zur phonetischen Ausgabe

Für die technische Realisierung von Sprachsynthese-Geräten lassen sich prinzipiell zwei unterschiedliche Verfahren nutzen:

— Reproduktive Synthese, d.h. Wiedergabe speziell codierter und gespeicherter Zeitfunktionsabschnitte natürlicher Sprache, womit jeweils nur ein begrenztes, bei der Datengenerierung vorgegebenes Vokabular synthetisierbar ist;

— Phonemgesteuerte Synthese, d.h. Bildung der Sprachsignale ausgehend von phonemähnlichen Bausteinen, so daß durch Ansteuerung eines derartigen Synthetisators mit einem phonetischen Code ein unbegrenztes Vokabular, ggf. sogar unterschiedlicher Sprachen, synthetisiert werden kann.

Als leistungsfähige Codierungsverfahren für die reproduktive Synthese haben sich das aus der Nachrichtentechnik bekannte Prinzip der adaptiven Delta-Puls-Code-Modulation (ADPCM), das LPC-Verfahren (Linear Predictive Coding) [2] sowie das vorrangig in Japan angewendete und mit LPC eng verwandte PARCOR-Verfahren (PARCOR partielle Autokorrelation) international durchgesetzt. Die im Stadium der Grundlagenforschung umfangreich untersuchte Formant-Synthese, bei der ein sehr niedriger Speicherplatzbedarf für die codierten Sprachdaten erforderlich ist (weniger als 1 kbit/s Sprache gegenüber 10 bis 20 kbit/s bei ADPCM und 2 bis 4 kbit/s bei LPC), hat nur geringe Anwendung in den auf dem Markt befindlichen Lösungen gefunden. Die Ursache hierfür ist darin zu sehen, daß die vorangehend genannten Verfahren sehr gut einer Realisierung durch LSI-Schaltkreise zugänglich sind. Es werden heute bereits mehr als 20 verschiedene Versionen für die reproduktive Synthese — vom Einchip-Sprachprozessor bis zum kompletten Gerät (vgl. Tafel 2) — sowie eine Vielzahl von Konsumgüter-Anwendungen angeboten. Da verfahrensbedingt mit ADPCM gegenüber LPC eine höhere Sprachqualität bei wesentlich geringerem Decodierungsaufwand, aber höherem Speicheraufwand erzielt werden kann, sind ADPCM-Lösungen insbesondere für die qualitativ hochwertige Synthese kleiner Vokabulare interessant, während LPC-Synthetisatoren bis zu einigen 100 Wörtern, bei Nutzung entsprechend großer Sprachdatenspeicher auch weit darüber hinaus, eingesetzt werden.

Für die phonemgesteuerte Synthese lassen sich vor allem zwei Lösungsmöglichkeiten hervorheben. Die eine beruht auf der rein künstlichen Nachbildung des menschlichen Spracherzeugungsprozesses, indem ausgehend von stimmhaften und stimmlosen Anregungssignalen mit Hilfe gesteuerter Filteranordnungen das Spektrum der einzelnen Laute sowie der Übergänge zwischen ihnen nachgebildet wird. Nach diesem Prinzip arbeiten

Tafel 2. Zusammenstellung der wichtigsten auf dem Markt befindlichen Sprachsynthese-Einrichtungen (nach Literaturangaben)

Typ	Verfahren	Sprachvolumen	Preis	Bemerkungen
UAA 1103	ADPCM	32 Wörter	30 DM	1 Chip einschl. 30 Kbit ROM
TSI S 2	ADPCM	64 Wörter	500 DM	3 Chips, davon 2 × 16 Kbit ROM
LPC-10	LPC	32 Wörter	?	1 Chip
HD 38 880	PARCOR	200 Wörter	?	3 Chips, darunter 1 × 128 Kbit ROM; bei Erweiterung auf 16 ROMs ≈ 30 min. Sprache
TMS 5100 TMS 5200	LPC	200 Wörter	150 DM	
MEA 8000	Formantsynthese reproduktiv	speicherabh. (8 KByte/min Sprache)	?	Synthese-Chip und Speicher
SYNTE 2	Formantsynthese phonemgesteuert	unbegrenzt	?	4 Leiterplatten
SC-01	Formantsynthese phonemgesteuert	unbegrenzt	12 $	1 Chip (Preis gilt für große Stückzahlen)
CDS II	Formantsynthese (Basis: SC-01)	unbegrenzt	15 000 $	Entw.-System mit Graphem-Phonem-Umsetzung
VSM	Formantsynthese (Basis: SC-01)	1300 Wörter 25 Suffixe und Präfixe	1200 $	Phonemcodes auf 8 KByte EPROM
VSB	Formantsynthese phonemgesteuert	unbegrenzt	280 $	OEM-Leiterplatte auf Basis SC-01

beispielsweise das gemeinsam vom VEB Robotron ZFT und der TU Dresden als Labormuster entwickelte Gerät ROSY 4200 oder das von der Firma Votrax Mitte der 70er Jahre auf den Markt gebrachte Gerät VS 6. Als neueste Lösung bietet Votrax mit dem SC-01 eine Einchip-Version für die komplette Synthese mit 64 Phonemen an. Auf der Basis dieses Schaltkreises wurden mehrere Synthesesysteme für unterschiedliche Anwendererfordernisse, bis hin zu einem Entwicklungssystem, das die Generierung der zur Ansteuerung des Synthetisators erforderlichen phonetischen Codes ausgehend von Buchstabentexten (Graphen-Phonem-Umwandlung) für Englisch übernimmt, entwickelt.

Eine zweite Möglichkeit für die phonemgesteuerte Synthese besteht darin, ausgehend von einer genügend großen Anzahl kontextbeeinflußter Laute (sog. Allophone) oder von Zweilautverbindungen (Diphone), die nach einem Verfahren der reproduktiven Synthese (z.B. LPC) gespeichert wurden, eine Synthese beliebiger Wörter durchzuführen. So wird in den USA auf der Basis eines LPC-Synthese-Chips auch eine für offenen Wortschatz geeignete Allophon-Synthese angeboten.

Der Vorteil des unbegrenzten Wortschatzes bzw. eines sehr geringen Speicherbedarfs für die Synthesecodes eines bestimmten Vokabulars (etwa 100 bit/s Sprache) bei der phonemgesteuerten Synthese wird durch eine im Vergleich zu den reproduktiven Verfahren wesentlich geringere Verständlichkeit und Natürlichkeit der erzeugten Sprache erreicht.

3. Sprach-Eingabe- und Ausgabegerät SEA 1600

Das als Forschungsmuster auf der Leipziger Frühjahrsmesse 1982 gezeigte Gerät SEA 1600 [4] enthält einen Spracherkennungsmodul SEM 1620 und einen Sprachsynthesemodul KAM (Kurzansagemodul). Beide Module sind gemeinsam mit den Stromversorgungsbaugruppen in einem Blockeinschub des Systems K 1600 untergebracht. Das Gerät SEA 1600 kann über ein Standard-Interface (z.B. 8-bit-Parallelinterface IFSP) als peripheres Gerät eines Rechners betrieben bzw. mit einem beliebigen anderen technischen System, das mit Sprachsteuerung versehen werden soll, gekoppelt werden. Die Spracheingabe

durch den Nutzer erfolgt über ein Nahbesprechungsmikrofon („Sprechgeschirr"). Für die Sprachausgabe ist ein extern anzuschließender Lautsprecher vorgesehen.
Der insgesamt 15 Steckeinheiten umfassende Spracherkennungsmodul arbeitet nach dem Prinzip der Einzelworterkennung mit einem Wortschatz bis zu etwa 100 Wörtern; die sprecherspezifische Referenzinformation wird durch einen Lernvorgang aus 3 bis 5 gesprochenen Realisierungen jedes Wortes ermittelt. Diese Referenzinformation wird abhängig vom jeweiligen Einsatzfall entweder von einem externen Speicher in den Operativspeicher des Erkennungsmoduls geladen oder im programmierbaren Festwertspeicher des Erkenners aufbewahrt. Bei einem Wortschatz von 72 Wörtern belegt die Referenzinformation eines Sprechers beispielsweise etwa 6 KByte. Als Merkmalvektoren werden beim SEM 1620 sog. Nulldurchgangs-Histogramme [3] benutzt. Das Entscheidungssystem arbeitet nach dem Referenzfolge-Verfahren mit dynamischer Optimierung. Die unter Laborbedingungen für ein Vokabular von 72 Wörtern ermittelte Erkennungssicherheit beträgt 97%. Die Reaktionszeit (Zeit zwischen Sprechbeginn und Anliegen des Erkennungsresultates am Interface) liegt maximal in der Größenordnung der Sprechdauer (einschließlich der die Wörter abschließenden Pausen), so daß unter Berücksichtigung dieser Pausen auch aufeinanderfolgend Informationen eingegeben werden können.
Der Sprachsynthesemodul KAM ist eine für die Erfordernisse des Fernmeldewesens entwickelte Sprachausgabeeinrichtung mit hoher Sprachqualität, die nach dem ADPCM-Verfahren arbeitet. Die auf den jeweiligen Anwendungsfall zugeschnittenen Sprachbausteine (Wörter, ggf. auch Silben) werden nach dem Datenkomprimierungsprozeß auf EPROM-Schaltkreisen gespeichert. Auf einer Steckeinheit mit 16 KByte Speichervolumen können Sprachbausteine mit einer Gesamtdauer von 4 s untergebracht werden. Anwendungsbeispiele haben gezeigt, daß auf Grund der freien Adressierbarkeit der Bausteine mit dem maximal verfügbaren Bausteinvolumen von 16 s bereits Texte von mehr als 1 min Gesamtdauer synthetisiert werden können.
Das System SEA 1600 wurde auf der Leipziger Frühjahrsmesse 1982 in Kopplung mit einem Rechner K 1630 in einem Demonstrationsbeispiel für die Sprachkommunikation eines Nutzers mit einem Lagerhaltungssystem vorgestellt. Eine ausführlichere Darstellung des Gesamtsystems wird in [4] gegeben.

Eingegangen am 15. September 1982 NaA 8988

Literatur

[1] *Bechen, P. v.:* Spracheingabe muß nicht teuer sein. Elektronik 30 (1981) 13, S. 52–54
[2] *Jaeger, G.:* Sprachdaten-Generierung nach dem Linear-Predictive-Coding-Verfahren. Nachrichtentechn., Elektron. 33 (1983) 1, S. 36–37
[3] *Kürbis, S.:* Nulldurchgangshistograph. Diplomarbeit, TU Dresden, Sektion Informationstechnik, 1981
[4] *Naumburger, V.; Püschel, K.:* SEA 1600 – ein Gerät zur phonetischen Ein- und Ausgabe. Neue Technik im Büro 26 (1982) 6, S. 161–163
[5] *Püschel, K.; Westendorf, Chr.-M.:* Das experimentelle System zur automatischen Einzelworterkennung EK III/2. Forschungsbericht Nr. 11/81, TU Dresden, Sektion Informationstechnik 1981
[6] *Weigert, G.:* Digitalfilterbank für die Sprachanalyse auf einem KRS 4201. Nachrichtentechn., Elektron. 33 (1983) 1, S. 30–32

Dr.-Ing. *Klaus Püschel*, VEB Robotron ZFT Fachbereich E 34, 8012 Dresden, Postschließfach 330

Psychoakustisches Funktionsmodell und Möglichkeiten einer technischen Systemoptimierung

F. Blutner, Zwota

Mitteilung aus dem VEB Forschungszentrum Musikinstrumente Plauen/Zwota

Deskriptoren: akustische Signalverarbeitung, Sprache, musikalische Akustik, physiologische Akustik

Sprache und Musik spielen als Ergebnis der Human-Evolution eine fundamentale Rolle in der menschlichen Gesellschaft. Mit der Entwicklung der Sprache hat die Natur einerseits die Möglichkeit geschaffen, Aussagen über die Wirklichkeit zu formulieren, mit den Kommunikationspartnern auszutauschen und den wissenschaftlichen Erkenntnisprozeß in Gang zu setzen. Demgegenüber ist die Musik in der Lage, die Emotionen unmittelbar und ohne Umwege anzusprechen, zu übertragen und somit sozial zugänglich zu machen.
Wissenschaft und Technik haben in ihrer langen Geschichte beiden Phänomenen große Aufmerksamkeit gewidmet. Bedeutende Leistungen des Musikinstrumentenbaus, der Raumakustik, Nachrichtentechnik, Elektroakustik u.a. zeugen von den praktischen Früchten dieses Prozesses und prägen entscheidend unsere moderne Zivilisation.
Mit der Entwicklung der Mikroelektronik entstehen gegenwärtig völlig neue Möglichkeiten der technischen Erzeugung und Verarbeitung von Musik und Sprache. Insbesondere betrifft das die Realisierung von Teilfunktionen (zunächst elementarer Natur), die bislang allein den Menschen vorbehalten waren, und es können bestimmte Prozeßabläufe dadurch rationeller und effektiver gestaltet werden, daß solche biologischen Teilfunktionen durch technische Teilfunktionen ersetzt werden (Bild 1).
Die intensive wissenschaftliche Auseinandersetzung mit dieser Problematik wird darüber hinaus eine Reihe neuer Wege z.B. zur medizinischen Behandlung von Defekten des Sprech- und Hörvermögens eröffnen, und es deutet sich schon heute die Möglichkeit an, das Leistungsvermögen individueller biologischer Teilprozesse mit technischen Systemen zu vervielfachen.

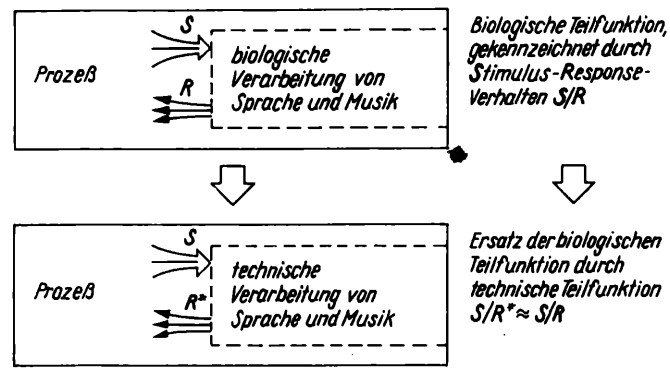


Bild 1. Zur Substitution biologischer Teilfunktionen der Verarbeitung von Sprache und Musik

1. Teilfunktionen der biologischen Verarbeitung von Sprache und Musik

Zur Lösung der in Bild 1 gezeigten Aufgabenstellung existieren heute eine Reihe von Forschungsansätzen. Die kybernetische Perzeptionsforschung orientiert sich dabei stark am biologischen Vorbild. Das Ziel besteht darin, ausgehend von Reiz-Reaktionsuntersuchungen ein Funktionsmodell der Empfindungs- und Wahrnehmungsvorgänge zu entwerfen. Dabei geht es in erster Linie um die Aufdeckung grundsätzlicher Funktionsprinzipien. Diese Prinzipien haben in der Evolution einen Selektionsvorteil verschafft und sind biologisch auf Grund der hohen Plastizität neuronaler Verknüpfungen weitgehend vollkommen realisiert. Die technische Nachbildung des biologischen Prinzips muß jedoch nicht die komplizierte Struktur neuronaler Verknüpfungen

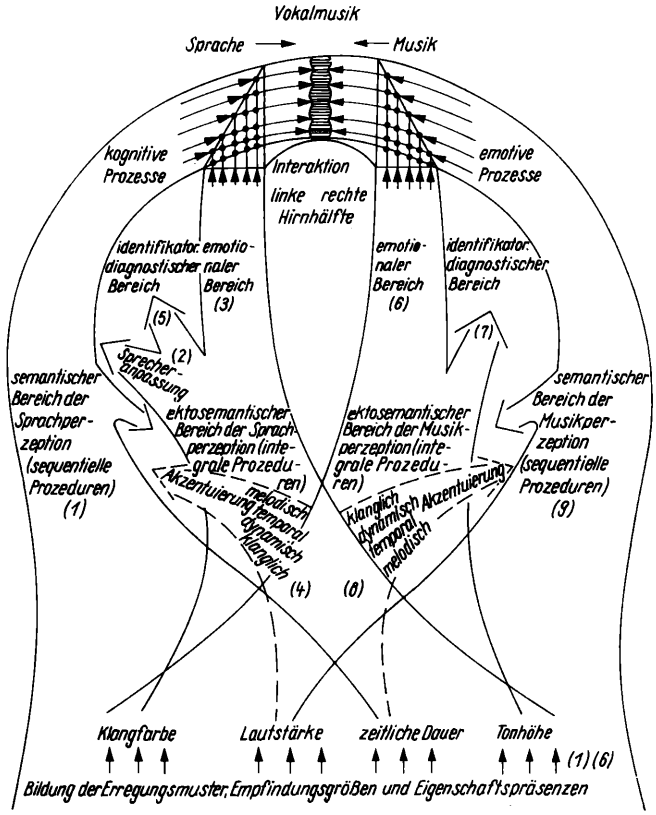


Bild 2. Überblick über die organische Musik- und Sprachverarbeitung

berücksichtigen. Vielmehr geht es um die technisch einfache Umsetzung von funktionalen Prinzipien entsprechend den flexiblen Möglichkeiten der modernen Mikroelektronik.

Aus biokybernetischer Sicht kann die humane Verarbeitung von Sprache und Musik als ein hochkomplexes System ineinandergreifender Mechanismen und spezialisierter Teilfunktionen charakterisiert werden. Bild 2 gibt entsprechend dem gegenwärtigen Kenntnisstand einen groben Überblick. Die Darstellung veranschaulicht, daß die gemeinsame Grundlage der organischen Sprach- und Musikverarbeitung durch die einheitlichen Empfindungsprozesse in den unteren Hierarchieebenen gegeben ist.

Die unterschiedlichen Funktionen von Sprache und Musik werden in den höheren Hierarchieebenen realisiert. Musik gehört zusammen mit bildnerischen Darstellungen zu den Zeichensystemen, die etwas „zeigen". Dabei werden über akustische Muster bestimmte emotionale Muster aufgezeigt. Demgegenüber gehört die Sprache zu den Systemen, die weitgehend auf konventionellen Regeln beruhen. Diese Gegenüberstellung von Sagen und Zeigen [1] hat offenbar ein unmittelbares anatomisches Korrelat.

Verschiedene Untersuchungen der Psychologie, Psychophysik und Hirnphysiologie erbrachten, daß die beiden Hemisphären des Gehirns in bemerkenswerter Arbeitsteilung zusammenwirken. Es zeigte sich, daß die sogenannte „dominante Hemisphäre" (in 97% aller Fälle ist das die linke Hirnhälfte) weitgehend auf die Prozesse der Sprachwahrnehmung und des verbalen Denkens spezialisiert ist, während die „untergeordnete Hemisphäre" vorzugsweise ganzheitliche Merkmale verarbeitet. Solche „holistischen" Eigenschaften können Formstrukturen der Musik oder bildnerischer Darstellungen sein, kurz „Zeigesysteme". Eine weitergehende Analyse der Funktion von Sprache und Musik zeigt das Zusammenwirken von zwei grundsätzlich verschiedenen Teilmechanismen (Bild 2):

1. Sequentielle Prozeduren des semantischen Bereichs
   Zuordnung bestimmter Signalmerkmale zu Zeichenklassen im Sinne der Semiotik.
2. Integrale Prozeduren des ektosemantischen Bereichs
   Zuordnung bestimmter Signalmerkmale zu i.allg. graduell variablen Eigenschaften des Expedienten (z.B. größere oder geringere Erregung eines Sprechers oder Musikers) und zur Identifizierung des Expedienten [2].

Die Prozeduren des semantischen und ektosemantischen Bereichs stehen in einem engen Zusammenhang (Bild 2). Musik und Sprache können auch in den Vokalgattungen über die im Bild gezeigten direkten Wege eng zusammenwirken. Darüber hinaus existieren noch eine Reihe indirekter Wirkungsmechanismen, die nur schwer zu formalisieren sind.

Die Abgrenzung von Teilfunktionen des biologischen Systems ist somit nur dann sinnvoll, wenn das im Kontext der komplexen Gesamtfunktion geschieht. Unter Beachtung dieser Tatsache ist es möglich, die im Bild 2 gezeigten Teilfunktionen im Rahmen bestimmter Handlungsabläufe durch äquivalente technische Teilfunktionen zu ersetzen. Bild 3 zeigt hierzu einige Beispiele aus der Sprachakustik und der musikalischen Akustik. Die Ziffern beziehen sich auf die in Bild 2 gekennzeichneten biologischen Teilfunktionen.

2. Psychoakustische Optimierung eines Akustikprozessors

Eine detaillierte Analyse der in Bild 3 zusammengestellten Anwendungsfälle zeigt, daß zur Realisierung der unterschiedlichsten Verarbeitungsfunktionen eine Auswahl von häufig wiederkehrenden Grundfunktionen ausreicht. Die gewünschten Teilfunktionen können somit nach dem Baukastenprinzip effektiv realisiert werden. Das erfordert folgende Schritte:

1. Bereitstellung von Elementarbausteinen mit hinreichend flexiblen Systemparametern
2. Auswahl der zur Lösung der Teilaufgabe benötigten Elementarbausteine und Spezifizierung der Systemparameter
3. Verknüpfung der Bausteine entsprechend der zu realisierenden Teilfunktion.

Diese grundsätzliche Vorgehensweise entspricht genau dem biologischen Evolutionsprozeß von Sprache und Musik [3]. Es bietet sich deshalb an, psychoakustische Erkenntnisse zur Optimierung von technischen Systembausteinen und Systemkomplexen der akustischen Informationsverarbeitung heranzuziehen.

Zur Erläuterung dieses Vorgehens soll nun ein einfacher Akustikprozessor entworfen werden, der vorzugsweise im Zeitbereich diskretisiert und somit besonders die dynamischen Signalparameter erfaßt.

Die erste Grundfunktion betrifft die Abbildung der Schalldruck-Zeitfunktion in einer für die nachfolgenden Verarbeitungsblöcke zweckmäßigen Form (Bild 4). Die Natur realisiert diese Funktion weitgehend durch periphere Informationsverarbeitungsprozesse im Innenohr und den unteren Schichten des Nucleus cochlearis. Diese Prinziplösung kann dabei als Reihenschaltung der in Bild 4 qualitativ angedeuteten Bausteine beschrieben werden [4].

Die technische Umsetzung dieses Prinzips wird entsprechend der geforderten hohen Zeitauflösung wie folgt vorgenommen:

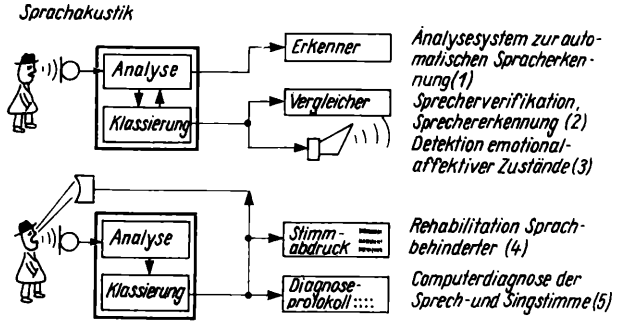
1. Grobe Frequenzselektion (etwa Oktav- bis Halboktavfilter).
2. Einweggleichrichtung der erhaltenen Signalfunktionen und Tiefpaßbegrenzung entsprechend der am ersten Neuron gemessenen Grenzfrequenz von 3 kHz.
3. Intensitätslogarithmierung.

Die resultierenden hüllkurvenverformten Frequenzkomponenten der Schalldruck-Zeitfunktion sind der Ausgangspunkt für die nachfolgende Stufe. Wie in Bild 5 am Beispiel einer vokalstimmhaften Konsonant-Vokal-Folge verdeutlicht, geht es nun darum, die in der transformierten Darstellung markant hervortretenden Signal-Periodizitäten und Hüllkurven-Verläufe zu selektieren.

Die Prinziplösung der ersten Teilaufgabe (A) basiert darauf, die abrupten Einsätze der gezeigten Periodenmuster mit Tiefpaßintegrator herauszufiltern [5]. Er muß in Anpassung an die zu verarbeitenden Signale eine kurze Aufladezeitkonstante zwischen 0,3 und 1 ms und eine etwa 10mal größere Entladezeitkonstante haben.

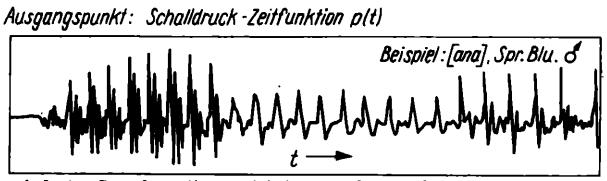
Die Prinziplösung der zweiten Teilaufgabe (B) basiert darauf, den Puls-Charakter der Signalverläufe durch den Effekt der Vor- und Nachverdeckung zu unterdrücken [6], gleichzeitig aber durch die spezifische Zeitcharakteristik der Verdeckungskurven die abrupten Hüllkurvenänderungen mit großer Zeitauflösung abzubilden.

Die technische Umsetzung des ersten Prinzips führte zur Ent-



Analysesystem zur automatischen Spracherkennung (1)
Sprecherverifikation, Sprechererkennung (2)
Detektion emotional-affektiver Zustände (3)
Rehabilitation Sprachbehinderter (4)
Computerdiagnose der Sprech- und Singstimme (5)
Steuerung technolog. Prozesse laufende Qualitätsüberwachung, Ursachenforschung (6)
Akustische Endkontrolle von Musikinstrumenten (7)
Kontrollgerät und Lernhilfe für die Musikerziehung (8)
Analysesystem für die Musikwissenschaft (9)

Bild 3. Applikationsbeispiele der technischen Verarbeitung von Sprache und Musik



Aufgabe: Transformation des Schalls in eine für nachfolgende Verarbeitungsprozesse zweckmäßige neuronale Erregungsstruktur

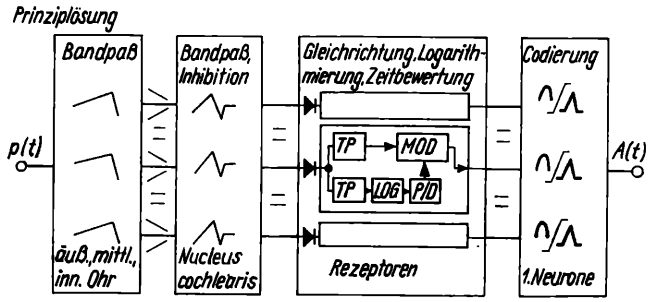
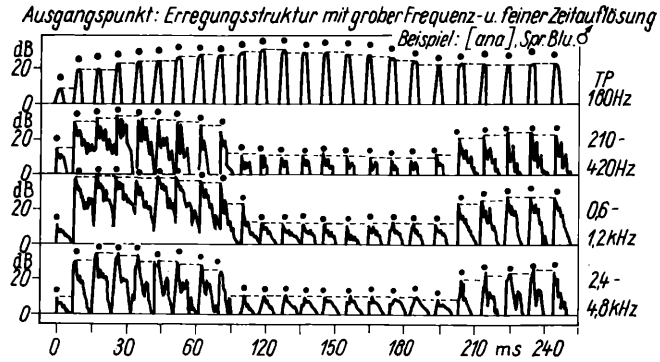


Bild 4. Funktionsprinzipien der primären Schallanalyse



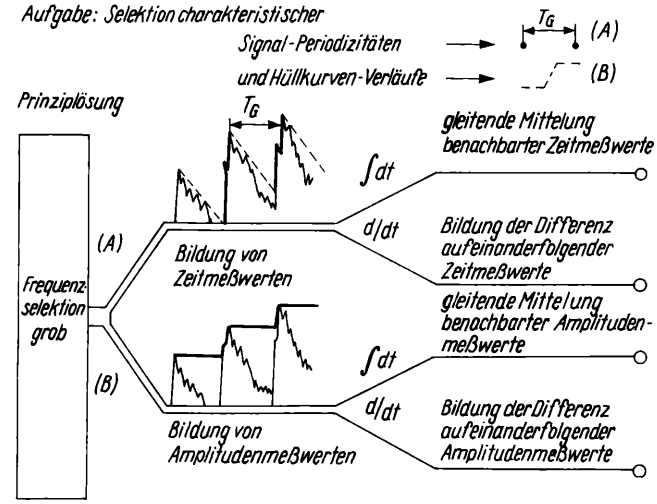
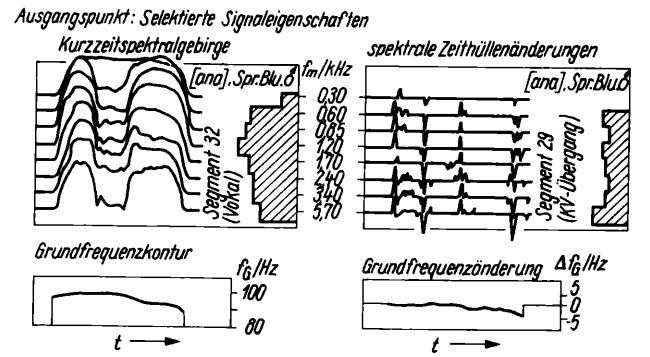


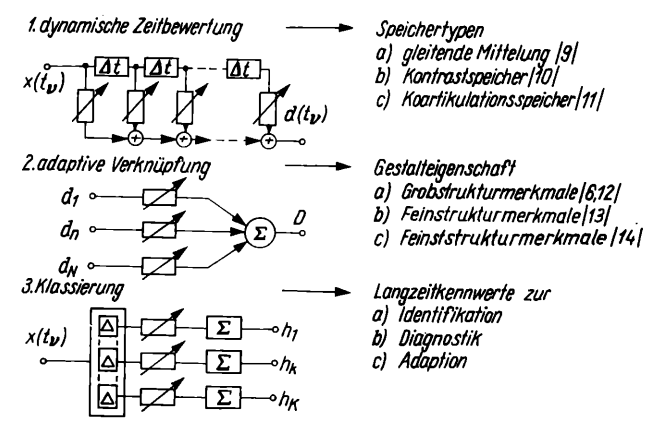
Bild 5. Funktionsprinzipien der akustischen Eigenschaftsselektion

Bild 6. Funktionsprinzipien der Verarbeitung akustischer Eigenschaftssätze



Aufgabe: Schrittweise Ableitung komplexer Gestaltparameter des zeitlich variablen Eigenschaftssatzes entsprechend einer vorgegebenen Disposition

Beispiele für elementare Teillösungen:



wicklung eines gesteuerten Tiefpaß-Integrators mit nachfolgender Impulsformung, Periodendauermessung und Fehlerkorrektur [7].

Die vorteilhafte technische Umsetzung des zweiten Prinzips ist mit getakteten Halteschaltungen und nachfolgender digitaler Meßwertverarbeitung möglich [8]. Die über gleitende Mittelung bzw. Differenzbildung weiter umgeformten Signalfunktionen sind der Ausgangspunkt der nächsten Etappe. Wie in Bild 6 dargestellt, geht es nun darum, bestimmte Gestaltqualitäten dieses hochdimensionalen Eigenschaftsatzes, entsprechend der angeborenen Disposition zur Sprach- und Musikverarbeitung, abzuleiten.

Die im Bild angedeuteten elementaren Prinziplösungen der Natur sind auch technisch vorteilhaft zu realisieren. Genannt seien in diesem Zusammenhang Perzeptron [15] und Lernmatrix [16], Klassiereinrichtungen [17] und Verfahren zur dynamischen

Folgentransformation [18]. Diese Systemlösungen haben in der Regel noch große Freiheitsgrade, die aber durch Berücksichtigung psychoakustischen Zusatzwissens sinnvoll eingeengt werden können.

Bild 7 zeigt zusammenfassend die Grobstruktur eines Mikrorechner-Experimentiersystems, das eine wahlweise Verknüpfung der oben beschriebenen Elementarfunktionen ermöglicht. Es ist vorgesehen, diese Systemkonfiguration als Pilotprojekt für verschiedene Anwendungen der Sprachakustik und musikalischen Akustik zu erproben.

Im nächsten Abschnitt sollen an Hand einfacher Beispiele einige Anwendungsmöglichkeiten veranschaulicht werden. Die Ergebnisse beziehen sich dabei i.allg. auf Simulationsexperimente an einem Kleinrechner.

3. Beispiele zur Erläuterung einiger Anwendungsmöglichkeiten

Als erstes Beispiel wird die Anwendung des Experimentiersystems zur Detektion einer lautsprachlichen Folgenstruktur erläutert. In Bild 8 wurden hierzu die Analyseergebnisse der von einem männlichen Sprecher realisierten Worte |Zander| und |Saite| dargestellt. Der Hüllkurvendemodulator wurde mit einer Zeitauflösung von 2,5 ms betrieben, wobei zur Ermittlung der Kurzzeitspektralgebirge jeweils 5 benachbarte Hüllkurvenwerte zusammengefaßt wurden. Die Darstellung zeigt zunächst eine sehr gute Abbildung auch kurzer Signalpausen, wie sie mit konventionellen Demodulatorschaltungen nicht erreicht wird. Das Prinzip der getakteten Hüllkurvendemodulation ermöglicht somit eine sehr sichere Pausendetektion.

Weiterhin werden neben den relativ stationären Signalabschnitten die abrupten Lautübergänge deutlich abgebildet. Zur genaueren Erfassung dieser Übergänge ist es zweckmäßig, die Differenzen benachbarter Hüllkurvenwerte im 2,5-ms-Raster zu bilden. Auf dieser Grundlage ist z.B. eine sehr sichere Unterscheidung der abrupt einsetzenden Affrikaten (wie in |Zander|) von den kontinuierlich einsetzenden Frikativlauten (wie in |Saite|) möglich (Erkennungsrate $> 95\,\%$). Weiterhin können Vokal-Nasallaut-Verbindungen (z.B. in |Zander|) segmentiert werden [10], während kontinuierliche Spektralveränderungen (z.B. Diphtong in |Saite|) keine Segmentmarken liefern.

Die in Bild 8 weiter dargestellten Ergebnisse der Grundfrequenzdetektion lassen zunächst den Charakter der Anregungsfunktion (stimmhaft/stimmlos) erkennen. Gemeinsam mit der Zeithülleninformation ist es damit z.B. möglich, eine sehr sichere Trennung der stimmhaften von den stimmlosen Explosivlauten vorzunehmen. So hat das [t] in |Saite| zwei um 40 ms deutlich versetzte On-Impulse und eine ebenfalls 40 ms lange Geräuschphase, während beim [d] in |Zander| der Abstand der Impulsmarken wesentlich kleiner ist und die Geräuschphase fehlt. Die Bedeutung des Verlaufs der Grundfrequenzkontur sei an einem nächsten Beispiel erläutert.

Bild 9 zeigt die experimentell ermittelten Grundfrequenzverläufe des von einem männlichen Sprecher mit unterschiedlicher Melodieführung gesprochenen Satzes |Das Wetter wird schön|. Dabei wurde eine Mittelung und ggf. Korrektur der Grundfrequenz über 3 Perioden vorgenommen und eine nichtlineare Zeitanpassung zur Elimination von Tempovariationen realisiert. Das Beispiel zeigt deutlich, wie durch die Melodieführung einerseits der Akzent gebildet wird (Anhebung der Grundfrequenz bei |Wetter|), zum anderen das Ende des Sprechaktes informativ gekennzeichnet und schließlich eine Emotion ausgedrückt wird.

Die Aneignung der richtigen Melodieführung ist ein wichtiges Problem der Rehabilitation Sprachbehinderter. Hierbei kann die geeignete optische Darstellung der Grundfrequenzkontur eine wesentliche Unterstützung leisten. Die maschinelle Auswertung des Grundfrequenzverlaufs liefert weiterhin wichtige Merkmale zur automatischen Sprechererkennung [19], zur Detektion emotional-affektiver Zustände und zur Ablaufsteuerung semantischer Auswahlprozesse (vgl. Bild 2).

Das nächste Beispiel erläutert die Anwendung des Verfahrens der Pegelklassierung zur Ermittlung statistischer Kennwerte fortlaufender Sprache. Bild 10 zeigt die Resultate der Klassierung einer etwa 10 s langen Sprachpassage. Die dick ausgezogene Kurve stellt das Langzeit-Mittelwertspektrum dar.

Die schraffierten Histogramme repräsentieren die Pegelverteilungen ausgewählter Filterkanäle (Frequenzgruppenfilterbank).

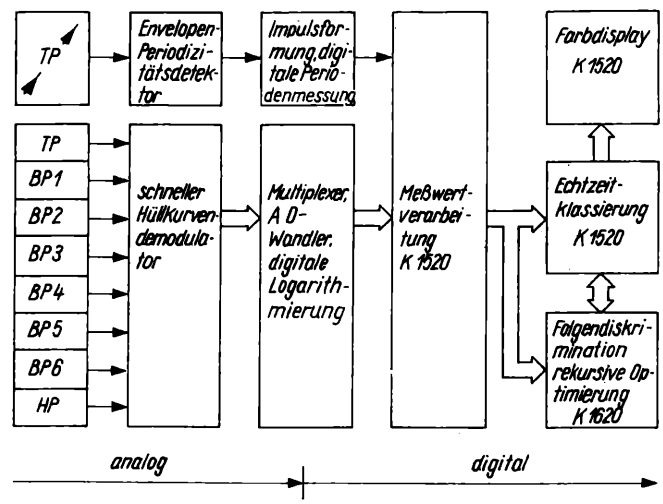


Bild 7. Blockschaltbild eines Akustikprozessors

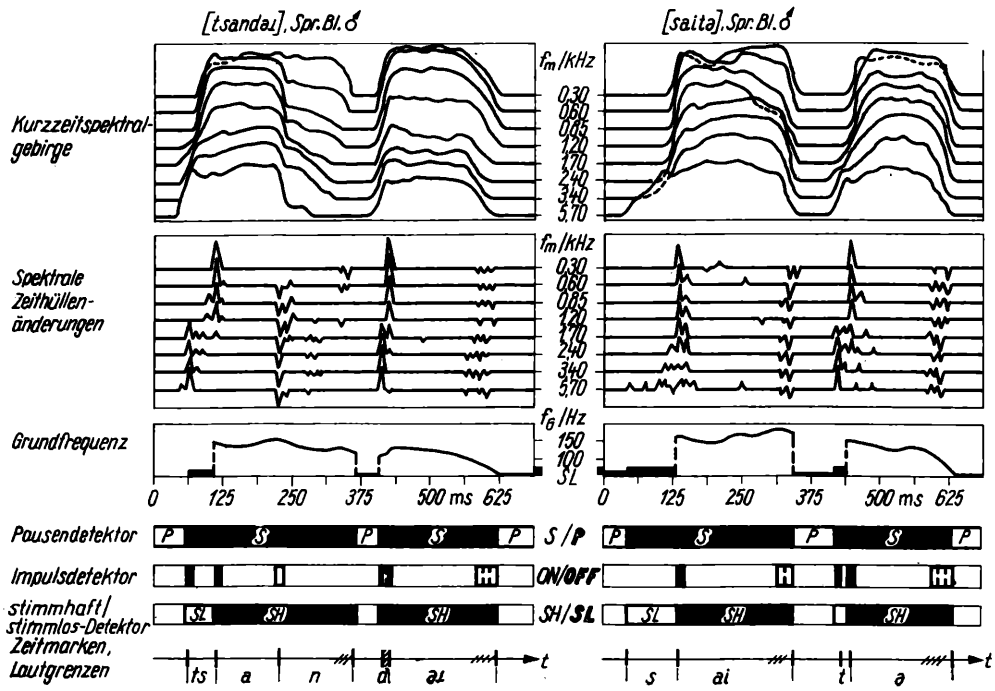
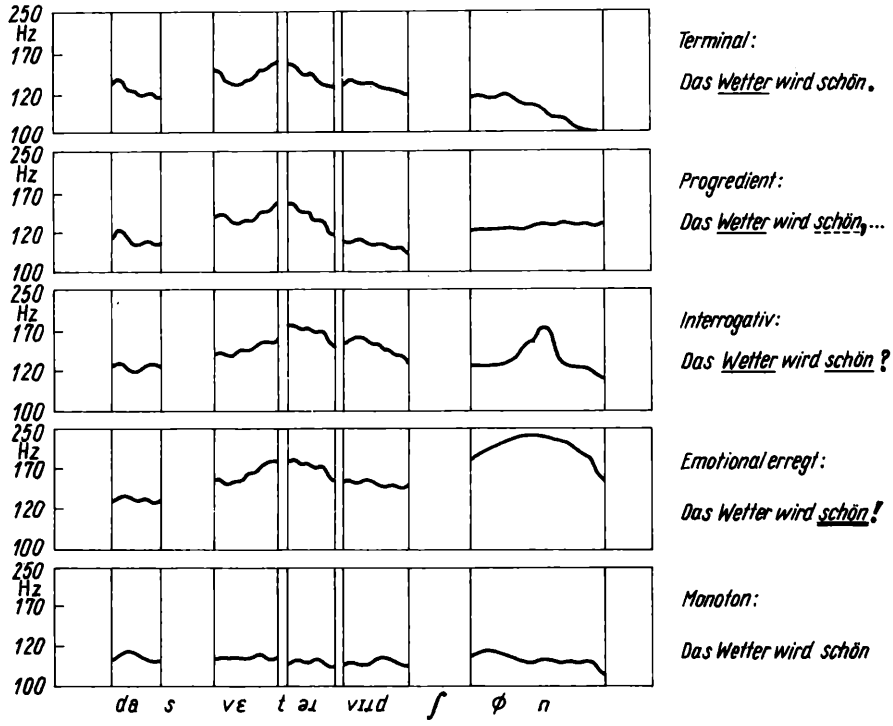


Bild 8. Beispiel zur Analyse einer lautsprachlichen Folgenstruktur

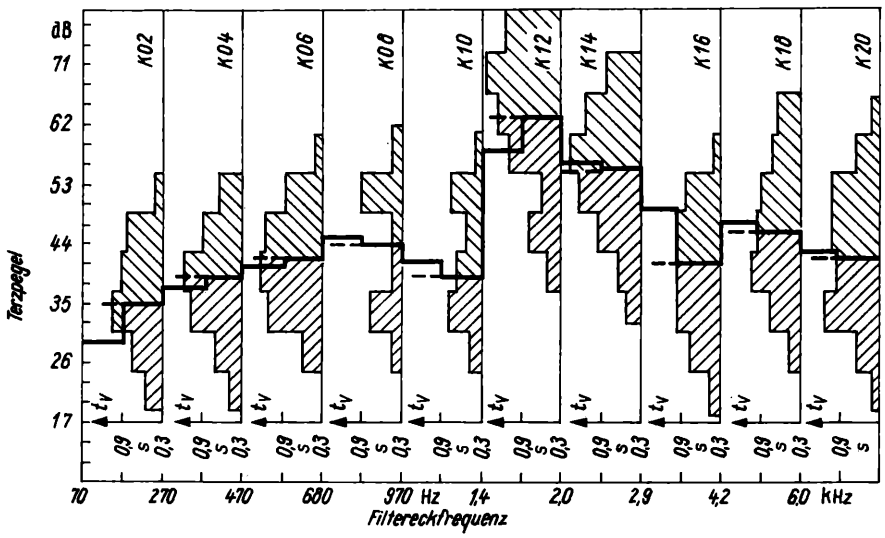
Bild 9. Beispiel zur Demonstration der unterschiedlichen Melodieführung in der Sprache

Bild 10. Pegelhistogramm fließender Sprache

Bild 11. Beispiel zur Analyse einer musikalischen Folgenstruktur

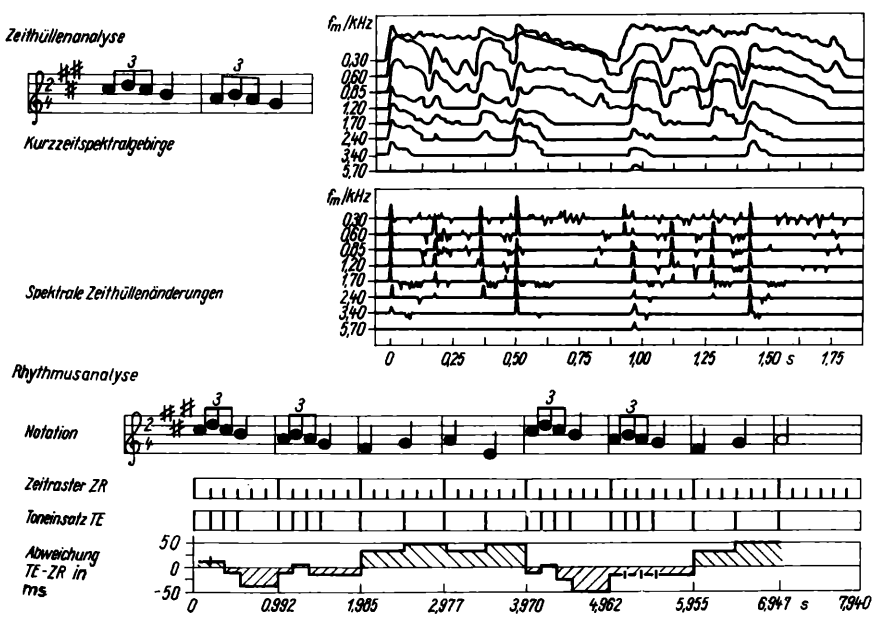


9



10

11



Das Langzeit-Mittelwertspektrum liefert grobe Merkmale zur Sprecherverifikation und zur Sprecheranpassung [20]. Die Bedeutung weiterer Verteilungsparameter der Pegelhistogramme ist noch nicht bekannt. Es gibt aber eine Reihe von Hinweisen dafür, daß wichtige identifikatorisch-diagnostische Merkmale abgeleitet werden können und eine effektive Steuerung dynamischer Klassifikationsprozesse zu realisieren ist.

Abschließend wird ein Beispiel aus der musikalischen Akustik erläutert. Bild 11 zeigt die Anwendung der Zeithüllenanalyse zur Ermittlung eines rhythmischen Musters. Analysiert wurde eine etwa 8 s lange, auf der Gitarre vorgetragene musikalische Phase. Wie im Bild dargestellt, können die Toneinsätze auf der Grundlage der im 2,5-ms-Zeitraster ermittelten Hüllkurvenänderungen exakt und sicher markiert werden. Diese Zeitmarken wurden nachfolgend zur Berechnung einer „kürzesten Notendauer" herangezogen. Damit ist das Zeitraster (Puls) der rhythmischen Folge bestimmt. Aus der Koinzidenz von Puls und gemessenen Zeitmarken kann nun auf die rhythmische Struktur selbst geschlossen werden. Weiterhin können (ggf. nach erneuter Zeitnormierung) systematische Abweichungen vom metrischen Schema berechnet werden. Die objektive Erfassung dieser Größen ist insbesondere für die Musikwissenschaft von großem Interesse, aber auch eine Hilfe für die Ausbildung professioneller Musiker.

Die in Bild 11 dargestellten Eigenschaftssätze liefern noch eine Reihe weiterer Informationen, insbesondere zur Klangqualität des Instruments. Untersuchungen zeigen, daß auf dieser Grundlage objektive akustische Gebrauchswertgutachten erstellt werden können, die dem kollektiven Urteil von 15 bis 25 Gutachtern gleichwertig sind [14].

4. Zusammenfassung

Es wurde gezeigt, daß psychoakustische Modellvorstellungen einen wichtigen Beitrag zur technischen Konzipierung leistungsfähiger akustischer Eigenschaftsselektionssysteme leisten können. Ausgangspunkt ist dabei ein allgemeines Perzeptionsmodell, das wesentliche funktionelle Prinzipien der organischen Sprach- und Musikverarbeitung in ihrer komplexen Wechselwirkung berücksichtigt. Hiervon abgeleitet, wurde die Funktion elementarer Modellkomponenten erläutert und die Konzeption eines leistungsfähigen Akustikprozessors begründet, der sich durch eine hohe Zeitauflösung auszeichnet.

Anwendungsfälle eines solchen universellen und flexiblen Systems sind insbesondere verschiedene Meßwertverarbeitungsaufgaben der Sprachakustik und der musikalischen Akustik. Das wurde abschließend an Beispielen aus beiden Teilbereichen näher erläutert.

Eingegangen am 15. September 1982 NaA 8989

Fortsetzung auf Seite 26

Clusterung zur Beurteilung von Analyseverfahren für die Sprachverarbeitung

U. Kordon, KDT, Dresden

Mitteilung aus der Technischen Universität Dresden, Sektion Informationstechnik

Deskriptoren: automatische Sprachverarbeitung, Signalanalyseverfahren, Bewertungsverfahren, Clusterung

1. Problemstellung

Bei der Konzeption von Sprachverarbeitungssystemen ist die Auswahl eines geeigneten Analyseverfahrens eine wesentliche Aufgabe. Während einerseits die technischen Möglichkeiten für die Realisierung und die angestrebten Leistungsparameter des Systems wesentlich für das zu verwendende Analyseverfahren sind, ist man andererseits bemüht, die Analyseeigenschaften der Verfahren bei typischen Signalkonfigurationen in die Einsatzentscheidung mit einzubeziehen. Prinzipiell muß festgestellt werden, daß das Hauptkriterium zur Beurteilung eines Analysesystems die Erkennungsraten des zugehörigen Klassifikators sind. Unter zugehörig soll verstanden werden, daß ein Analysesystem nicht unabhängig beurteilt werden kann, sondern daß stets das Gesamtsystem zu sehen ist, für das dieses Analysesystem verwendet werden soll. Jedoch hat die alleinige Anwendung der Erkennungsrate als Bewertungsbasis für den Analysator folgende Nachteile:

1. Bei Voraussetzung des Worterkennungsprinzips ist keine eindeutige Lokalisierung des eventuell fehlerhaft analysierten Signalabschnitts mehr möglich.
2. Die Aussagen sind sehr stark von den verwendeten Erkennungseinheiten abhängig (d.h. auch von eventuellen Segmentierungsverfahren) und bringen neben der Leistungsfähigkeit des Analysators in gleichem Maß die des Klassifikators mit zum Ausdruck.

Hinzu kommt, daß es oftmals erforderlich ist, bereits vor der Realisierung des Klassifikationsteils Informationen über bestimmte Eigenschaften des Analysesystems zu ermitteln, um diese für die Konzeption des Klassifikators zu nutzen. Die Kriterien für die Beurteilung von Analyseverfahren für Sprachverarbeitungssysteme können damit folgendermaßen zusammengefaßt werden:

a) Erkennungsrate des zugehörigen Klassifikators
b) Welche Abschnitte des Sprachsignals werden durch das jeweilige Analyseverfahren unterschieden?
c) Welches Verhalten haben die Analyseverfahren bei der Abbildung verschiedener typischer Signaleigenschaften?
d) Welcher Aufwand ist für die Anwendung der Analyseverfahren erforderlich?

Während das Vorgehen zur Erlangung von Aussagen zu a) und d) klar ist, soll im vorliegenden Beitrag auf Möglichkeiten zur Erlangung von Aussagen zu Kriterium b) und c) eingegangen werden.

2. Lösungsmöglichkeiten

Die bisherigen Lösungsansätze zur Ermittlung von Angaben zu b) und c) können in zwei Gruppen eingeteilt werden.

2.1. Bewertung durch Gütemaße der Strukturierung des Merkmalraumes

Dieses Vorgehen wurde vor allem gewählt, um den starken Einfluß des Klassifikators auf das Bewertungsergebnis zu umgehen. Grundgedanke ist die Formulierung von geeigneten Gütemaßen, die die Geometrie und Strukturierung des Merkmalraumes für das jeweilige Analyseverfahren zum Ausdruck bringen soll (z.B. Kompaktheit 1). Diese Maße geben praktisch an, wie stark das Analyseergebnis für ähnliche Signalkonfigurationen streut bzw. in welchem Maß unterschiedliche Signalkonfigurationen in einer deutlich unterschiedlichen Beschreibung durch das Analyseverfahren zum Ausdruck kommen.
Wenn man statt Abschnitten mit bestimmter Signalkonfiguration die eigentlichen Erkennungseinheiten nutzt (z.B. Phoneme), lassen sich Vorhersagen treffen, wie gut sich diese Klassen durch einen Klassifikator trennen lassen. Bei einer Einschätzung solcher Maße muß festgestellt werden, daß zwar Aussagen zu b) und c) getroffen werden können, jedoch hängen diese sehr stark von den Abschnitten ab, auf deren Basis sie formuliert werden.

2.2. Bewertung an Hand der Erfüllung bestimmter Vorgabekriterien

Wenn bei der Konzeption des Analyseteils bereits weitestgehende Klarheit über Art und Dimension des angewendeten Merkmalsatzes besteht, kann auf eine möglichst allgemeine Beurteilung verzichtet werden und als Kriterium die Genauigkeit der Wider-

Fortsetzung von Seite 25

Literatur

[1] *Bierwisch, M.:* Musik und Sprache. Überlegungen zu ihrer Struktur und Funktionsweise. Jahrbuch Peters 1978. *E. Klemm* Leipzig: Edition Peters 1979, S. 9–102
[2] *Meier-Eppler, W.:* Grundlagen und Anwendungen der Informationstheorie. 2. Aufl. Berlin, Göttingen, Heidelberg: Springer-Verlag 1969
[3] *Höpp, G.:* Evolution der Sprache und Vernunft. Berlin, Heidelberg, New York: Springer-Verlag 1970
[4] *Schönfeld, M.:* Die Verarbeitung akustischer Signale zwischen äußerem Ohr und 1. akustischen Neutronen. Dissertation, TU Dresden, 1974
[5] *Filip, M.:* Envelope periodicity detection. J. Acoust. Soc. of Am. 45 (1969) S. 719–732
[6] *Zwicker, E.; Feldtkeller, R.:* Das Ohr als Nachrichtenempfänger. 2. Aufl. Stuttgart: S. Hirzel-Verlag 1967
[7] *Blutner, F.; Fleischer, P.:* Verfahren und Anordnung zur Bestimmung der Grundperioden von Sprach- und Musiksignalen (in Vorbereitung)
[8] *Blutner, F.; Müller, W.; Hausfeld, H.:* Verfahren und Anordnung zur schnellen Hüllkurvenmodulation von Sprach- und Musiksignalen (in Vorbereitung)
[9] *Zwicker, E.:* A model describing temporal effects in loudness and threshold. Papers 6 th. Intern. Congress on Acoustics, Tokyo 1968, A-3-4
[10] *Hausfeld, H.:* Vokal-Konsonant-Unterscheidung unter Ausnutzung spezieller Hüllkurveninformationen. Vortrag zum 16. Fachkoll. Informationstechnik, Dresden 1983
[11] *Blutner, F.:* Modell zur organischen Sprachverarbeitung. Dissertation TU Dresden 1975
[12] *Bismarck, G. v.:* Extraktion und Messung von Merkmalen der Klangfarbenwahrnehmung stationärer Schalle. Dissertation TU München 1972
[13] *Adam, N.:* Prinzipien bei der subjektiven Klassifikation von Vokalen. Dissertation, TU Dresden, 1974
[14] *Haller, K.; Blutner, F.:* Multiples Modell zur Beschreibung subjektiver Klangattribute. Vortrag zum 16. Fachkoll. Informationstechnik, Dresden 1983
[15] *Rosenblatt, F.:* The perceptron, a probalistic model for information storage and organisation in the brain. Psychol. Rev. 65 (1958) S. 386–408
[16] *Steinbuch, K.:* Lernmatrix. Kybernetik 1 (1961) 1, S. 36–45
[17] *Zander, H.:* Automatisierte Pegelklassierung. Internat. Elektronische Rundschau 6 (1974) 6, S. 117–119
[18] *Weigert, G.:* Verfahren zur dynamischen Transformation von Merkmal-Vektorfolgen (persönl. Inform.)
[19] *Ney, H.:* Bestimmung der Zeitverläufe von Intensität und Grundperiode der Sprache für die automatische Sprechererkennung. Frequenz 35 (1981) 10, S. 265–270
[20] *Brouwer, P.:* Ein Spracherkennungsgerät mit automatischer Sprecheranpassung. Frequenz 32 (1978) 7, S. 204–207

Dr.-Ing. *Friedrich Blutner*, Institut für Musikinstrumentenbau, 9657 Zwota (Vogtland), Klingenthaler Str. 42

spiegelung bestimmter einzelner Signalkenngrößen im Analyseergebnis herangezogen werden (z.B. Grundfrequenz [2], Formantfrequenzen [3]). Dieses Vorgehen ist nicht unbedingt an einen klassifizierbaren Signalabschnitt gebunden, so daß eine Beeinflussung des Ergebnisses durch ein Segmentierungsverfahren ausgeschlossen wird. Jedoch können mit dieser Methode nur ähnliche Analyseverfahren verglichen werden. Ein weiterer Nachteil besteht darin, daß ein Sollwert erforderlich ist. Zur Bestimmung dieses Vergleichswerts ist aber in der Mehrzahl der Fälle bereits selbst ein meist nur Näherungen liefernder Analyseprozeß erforderlich. Im Sinne der eingangs formulierten Kriterien wird vor allem die Frage der selektiven Abbildungseigenschaften bei unterschiedlichen Signalkonfigurationen der Analyseverfahren nur ungenügend berücksichtigt. Im folgenden Abschnitt wird deshalb die Anwendung einer automatischen Strukturierung auf Sprachsignale vorgestellt, auf deren Basis Beurteilungsgrößen für Analyseverfahren definiert werden, die die angeführten Nachteile der oben diskutierten Methoden nicht mehr haben.

3. Clusterverfahren

Wie bereits erläutert, ist ein wesentlicher Nachteil der bisher gebräuchlichen Bewertungsmethoden für Analyseverfahren die fast immer erforderliche Segmentierung des kontinuierlichen Sprachsignals in geeignete Abschnitte. Eine günstige Lösung dieses Problems wäre die automatische Strukturierung des Sprachsignals, d.h., es erfolgt keine vorherige Definition von Signalabschnitten, deren differenzierte und reproduzierbare Beschreibung Bewertungsgrundlage für das Analyseverfahren darstellt, sondern die Abschnitte werden auf der Basis von Ähnlichkeitskriterien durch Zusammenfassen von Merkmalvektoren der das zeitlich begrenzte Sprachsignal beschreibenden und vom jeweiligen Analyseverfahren gelieferten Folge von Merkmalvektoren für kurze äquidistante Abschnitte konstanter Länge gebildet. Die dabei entstehenden Segmente entsprechen somit direkt bestimmten Signaleigenschaften, die durch das jeweilige Verfahren differenziert abgebildet werden können. Anzahl, Größe, Struktur und Lage zum Signal der entstehenden Segmente ermöglichen es, gezielte Aussagen über wesentliche Eigenschaften des jeweiligen Analyseverfahrens zu treffen.

Aus der Vielzahl der zur Ermittlung derartiger Strukturierungen bekannten Clusterverfahren [4] [5] wurde ein mehrstufiges serielles Verfahren ausgewählt. Die als Basisinformation dienende Merkmalvektorfolge ist $X = \{\vec{x}_t\} = (\vec{x}_1\ \vec{x}_2\ \ldots\ \vec{x}_T)$, wobei $\vec{x}_1$ der Merkmalvektor für den ersten Signalabschnitt konstanter Länge, $\vec{x}_2$ für den nachfolgenden usw. ist. T ist damit die Gesamtzahl dieser Abschnitte. Die durch Zusammenfassung mehrerer $\vec{x}_t$ entstehenden größeren Segmente C_i sollen jeweils durch den Mittelwert $\vec{\mu}_t$ der zu einem solchen Cluster C zusammengefaßten $\vec{x}$ repräsentiert werden. Schließlich wird die Entscheidung, daß ein Vektor $\vec{x}_t$ der Folge X zu einem bereits existierenden Cluster C zugeordnet wird, dann gefällt, wenn ein Abstandswert g zwischen dem jeweiligen $\vec{x}_t$ und $\vec{\mu}_i$ eine vorgebbare Schwelle g_s unterschreitet. Für g wurde hier der euklidische Abstand gewählt, jedoch sind prinzipiell alle aus der Objekterkennungstheorie bekannten Distanzmaße anwendbar.

In einer ersten Stufe wird zunächst eine Anfangsstrukturierung nach folgendem Algorithmus bestimmt:

— Startcluster $C_1\ \vec{\mu}_1 = \vec{x}_1,\ w = 1$

— Zuordnung $x_t \in C_i \mid i = 1, \ldots, w$

(für $t = 2, \ldots, T$) $\mid\ g(\vec{x}_t/C_i, \vec{\mu}_i) \leqq g_s$

falls $\vec{x}_t \notin C_i \rightarrow w = w + 1,\ \vec{\mu}_w = \vec{x}_t$.

Es ist ersichtlich, daß alle Vektoren streng seriell verarbeitet werden. Damit ist es aber möglich, daß ein Vektor $\vec{x}_t$ zu einem Cluster C_i gehören müßte, das aber erst nach seiner Verarbeitung gebildet wurde. Deshalb wird in einer weiteren Stufe eine

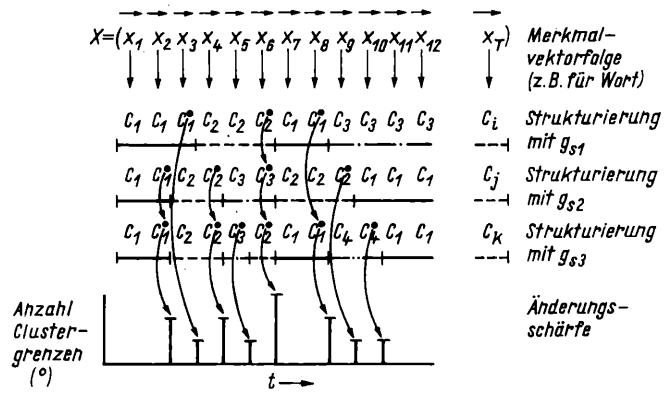


Bild 1. Zur Bildung des Auswertekriteriums Änderungsschärfe (willkürlich gewähltes qualitatives Beispiel)

nochmalige Zuordnung aller $\vec{x}$ zur nunmehr vollständig vorliegenden Clustermenge vorgenommen. Außerdem wird von der nichtdisjunkten Zuordnung zu einer disjunkten übergegangen, d.h. jeder Vektor $\vec{x}_t$ wird nur dem Cluster zugeordnet, zu dem g minimal ist, d.h. für

$$t = 1, \ldots, T, i = 1, \ldots, w: \vec{x}_t \in C_i \mid g(\vec{x}_t/C_i, \vec{\mu}_i) \rightarrow \text{Min.}$$

In der letzten Stufe erfolgt nun die Schätzung der clusterspezifischen Mittelwerte $\vec{\mu}_i$ als Mittelwert aller zum jeweiligen Cluster C_i zugeordneten $\vec{x}$ und eine erneute Zuordnung aller $\vec{x}_t$ nach minimalem g zu diesen korrigierten Clustern. Diese Stufe wird wiederholt, bis alle Cluster bezüglich $\vec{\mu}_i$ und Belegung konstant bleiben. Mit dem Schwellwert g_s ist praktisch die Feinheit der zu ermittelnden Strukturierung steuerbar.

4. Auswertekriterien, Ergebnisse

Zur Auswertung der nun vorliegenden Strukturierungen werden zwei Auswertekriterien formuliert.

Tafel 1. Für verschiedene Analyseverfahren ermittelte Anzahl von Übergängen zwischen den Lauten einer ausgewählten Lautgruppe zu den jeweils übrigen ausgewählten Lautgruppen, die in mindestens zwei Strukturierungen derselben Sprachprobe mit unterschiedlichem Schwellwert auftreten, bezogen auf die Gesamtzahl der in der verwendeten Sprachprobe insgesamt enthaltenen Übergänge zwischen den entsprechenden Lautgruppen

	Explosivlaute		Diphtonge	Nasale	Vokale
	stimmhaft	stimmlos			
Fast-Fourier	0,395	0,345	0,855	0,735	0,815
Nulldurchgang	0,33	0,4	1,085	0,87	0,515
Linear-Prediction	0,28	0,33	0,835	0,845	0,72

Tafel 2. Anzahl der in mindestens zwei Strukturierungen derselben Sprachprobe (mit unterschiedlichem Schwellwert) auftretenden Clustergrenzen innerhalb von verschiedenen Lautgruppen zuordenbaren Signalabschnitten (gemittelt über alle in der verwendeten Sprachprobe vorkommenden entsprechenden Signalabschnitte)

	Explosivlaute		Diphtonge	Nasale	Vokale
	stimmhaft	stimmlos			
Fast-Fourier	1,39	2,15	2,00	1,21	2,20
Nulldurchgang	0,88	1,62	2,38	2,53	2,12
Linear-Prediction	0,78	1,04	3,88	2,15	3,08

Tafel 3. Anzahl der durch Clusterung aller Spracheinheiten der verwendeten Sprachprobe ermittelten Grundsegmente, zu denen Laute ausgewählter Lautgruppen zugeordnet wurden, bezogen auf die Gesamtzahl der ermittelten Grundsegmente

	Explosivlaute		Diphtonge	Nasale	Vokale
	stimmhaft	stimmlos			
Fast-Fourier	0	0	0,203	0,04	0,395
Nulldurchgang	0,21	0,17	0,13	0,42	0,46
Linear-Prediction	0,04	0,23	0,206	0,31	0,575

4.1. Auswertekriterium Änderungsschärfe

Diese Größe soll angeben, wie deutlich Änderungen der Signalkonfiguration im Analyseergebnis zum Ausdruck kommen. Dazu wird das Clusterverfahren mehrmals mit verschiedenen Schwellwerten auf die entsprechende Spracheinheit (z.B. Worte) angewandt. Als Änderungsschärfe wird nun über der Zeit aufgetragen, in wieviel Strukturierungen zu dem jeweiligen Zeitpunkt eine Clustergrenze vorliegt. Eine durch das entsprechende Analyseverfahren erfaßte Änderung der Signalkonfiguration wird sich also in allen Strukturierungen in Form einer oder mehrerer Clustergrenzen niederschlagen, während weniger ausgeprägte nur bei einer hohen Schwelle zu einer Clustergrenze führen (Bild 1).

4.2. Auswertekriterium Grundsegmentsysteme

Um zu ermitteln, welche Grundsignalkonfigurationen unterschieden werden, wird das Clusterverfahren auf alle zu untersuchenden Spracheinheiten angewandt, so daß die entstehenden Cluster bei hinreichend großer Sprachprobe als die durch das entsprechende Analyseverfahren sicher unterschiedlich abgebildeten Signaleigenschaften interpretiert werden können.
In den Tafeln 1 bis 3 sind auf diesen beiden Auswertekriterien basierende Ergebnisse für drei verschiedene Analyseverfahren (Linear-Prediction-Analyse, 15 Komponenten, stationäre Methode; Fast-Fourier-Analyse, in 17 Kanäle komprimiertes Betragsspektrum; Nulldurchgangsanalyse, in 16 Kanäle komprimiertes Nulldurchgangshistogramm) angegeben. Als Sprachmaterial wurden 52 Worte von 2 Sprechern verwendet. Das Zeitfenster (*Hamming*) zur Ermittlung der Vektoren der Basisvektorfolge X betrug 25 ms, jeweils um 10 ms fortgesetzt. Zur quantitativen Erfassung des Beurteilungskriteriums b) aus Abschn. 1 wurde ein Trennmaß zwischen verschiedenen ausgewählten Signalkonfigurationen (Lautgruppen) definiert. Es entspricht dem Verhältnis zwischen Anzahl der Lautgruppenübergänge von einer Lautgruppe zu den jeweils übrigen Lautgruppen, für die die Änderungsschärfe einen vorgebbaren Mindestwert (hier 2) übersteigt und der Anzahl der in der Sprachprobe insgesamt vorhandenen Übergänge zwischen dieser Lautgruppe und den übrigen. Es ist damit unmittelbar ein Maß für das differenzierte Abbildungsvermögen dieser Signalkonfigurationen durch das Analyseverfahren (Tafel 1). Die Anzahl der Clustergrenzen innerhalb einer bestimmten Signalkonfiguration (Lautgruppe) soll schließlich zur quantitativen Formulierung von Beurteilungskriterium c) herangezogen werden (Tafel 2). Je weniger Clustergrenzen innerhalb eines stationären Signalabschnitts liegen, desto stabiler, d.h. eindeutiger, ist die Beschreibung dieser Abschnitte, während eine günstige Beschreibung dynamischer Signalphasen (z.B. Diphtonge) durch mehrere Clustergrenzen innerhalb dieser Segmente zum Ausdruck kommt. Diese Werte sollen abschließend durch die Verteilung einiger Lautgruppen auf die Grundsegmente ergänzt werden. Tafel 3 gibt an, wieviel Prozent der ermittelten Grundsegmente (≈ 25) vokalische, nasale usw. Abschnitte enthalten. Ein hoher Wert deutet dabei darauf hin, daß bereits geringe Unterschiede in den Signalen der Laute der jeweiligen Lautgruppe im Analyseergebnis erfaßt werden, also das jeweilige Analyseverfahren besonders selektiv für diese Signalkonfiguration ist.

5. Schlußbemerkungen

Diese hier vorgeschlagenen Größen zur Beurteilung von Analyseverfahren sind als Ergänzung der Erkennungsrate und des Aufwandes als Kriterium zu verstehen, um vor allem bei Spracherkennungssystemen auf Wortebene genauere Informationen über die Ursache möglicher Fehlerkennungen zu erlangen. Dabei liegt der eigentliche Vorteil bei der Untersuchung einzelner Spracheinheiten, um Aufschluß über die Güte der durch verschiedene Analyseverfahren gelieferten Beschreibungen dieser Einheiten zu erlangen. Es ist verständlich, daß die angegebenen Werte von der jeweiligen Spracheinheit und dem Schwellwert g_s abhängen. Somit müssen die hier angegebenen globalen Ergebnisse vor allem unter dem Aspekt des Relativvergleichs zwischen den drei Analyseverfahren gesehen werden. Trotzdem muß von diesen Ergebnissen abgeleitet werden, daß die Unterschiede dieser derzeit gebräuchlichsten Analyseverfahren bezüglich Kriterium b) bzw. c) nur gering sind, so daß das Hauptaugenmerk vordergründig auf eine optimale (dynamische) Anpassung an die jeweilige Signalkonfiguration zu richten ist. Eine ausführliche Diskussion zu der in diesem Beitrag vorgestellten Bewertungsmethode enthält [6].

Eingegangen am 15. September 1982 NaA 8090

Literatur

[1] *Fuchs, S.:* Ein methodischer Beitrag zur Lösung des Problems der automatischen Spracherkennung. Dissertation A, TU Dresden, 1969
[2] *Rabiner, L. R.*, u.a.: A comparative performance study of several pitch detection algorithms. IEEE Trans. on acoustics, speech and signalprocessing 24 (1976) 5, S. 399 – 418
[3] *Keller, T. e.:* Die Kennzeichnung von Sprachlauten durch Spektrum, Autokorrelationsfunktion und Nulldurchgangsabstände. Dissertation A, TH Karlsruhe, 1966
[4] *Peipmann, R.:* Grundlagen der technischen Erkennung. Berlin: VEB Verlag Technik 1975
[5] *Bock, H. H.:* Automatische Klassifikation. Göttingen: Vandenhoeck & Ruprecht 1974
[6] *Kordon, U.:* Eigenschaften von digitalen Analyseverfahren für die Sprachverarbeitung. Dissertation A, TU Dresden, 1981

Dr.-Ing. *Ulrich Kordon*, Technische Universität Dresden, Sektion 9 Informationstechnik, 8027 Dresden, Mommsenstr. 13

Verfahren zur dynamischen Sprecheranpassung

K.-H. Affa, Dresden

Mitteilung aus der Technischen Universität Dresden, Sektion Informationstechnik

Deskriptoren: Automatische Spracherkennung, Sprechererkennung, Sprecheranpassung, Sprecheranpassung adaptives System

1. Ausgangspunkt

Mit der Einführung der Sprache in den Kommunikationsprozeß zwischen Mensch und Maschine ergeben sich Möglichkeiten, den Informationsaustausch noch effektiver zu gestalten. Für die Sprachausgabe existieren bereits eine Reihe von Systemen, die menschliche Sprache in verständlicher Form artikulieren können. Weitaus schwieriger ist die Problematik der automatischen Erkennung von Sprache. Ein grundlegendes Problem ist dabei die Individualität des Sprachsignals. Es ist außerordentlich schwierig, die informationstragenden Merkmale der Spracheinheiten zu extrahieren und damit die Beeinflussung einer lautlichen Äußerung durch unterschiedliche Sprecher und unterschiedliche Sprechbedingungen weitgehend zu kompensieren. Aus diesem Grund sind bei der Anwendung von automatischen Spracherkennungssystemen Einschränkungen nötig. Entweder man reduziert die zur Erkennung herangezogene Wortzahl bzw. die Anzahl der möglichen Sprecher oder man versucht durch Einbeziehung syntaktischer und semantischer Korrekturen Erkennungsfehler zu kompensieren [1]. Dabei konnte die Berücksichtigung solcher Korrekturen in technischen Systemen auf Grund der Komplexität dieses Problems nur in beschränktem Umfang erfolgen. Sie sind bisher nur für den Laborfall von Interesse. Gefragt sind jedoch möglichst solche Lösungen, die für einen größeren Anwenderkreis bei vertretbarem Aufwand dienen können.

2. Möglichkeiten zur Reduzierung des Sprechereinflusses

Je nach Anwendungsfall lassen sich unterschiedliche Verfahren zur Reduzierung des Sprechereinflusses einsetzen. Am einfachsten ist es, wenn das Erkennungssystem nur für einen einzigen Sprecher ausgelegt ist. Immer wenn ein neuer Sprecher das System benutzen will, muß neu angelernt bzw. ein neuer Referenzdatensatz geladen werden. In diesem Fall wird also der diesem Sprecher zugeordnete Datensatz zur Erkennung herangezogen. Von einer eigentlichen Reduzierung des Sprechereinflusses kann in diesem Zusammenhang jedoch noch nicht gesprochen werden, da jeder Sprecherwechsel einen neuen Referenzdatensatz nötig macht.

Andere Systeme arbeiten mit einem Referenzdatensatz, der von einer Vielzahl von Sprechern gebildet wurde. Somit existiert für jedes zu erkennende Wort ein Satz von Referenzmustern, mit genau so viel Einzelworten, wie Sprecher berücksichtigt wurden. Im Gegensatz zum oben erwähnten Verfahren wird hierbei ein Referenzdatensatz zur Erkennung herangezogen, der um ein Vielfaches umfangreicher ist. Das erfordert nicht nur einen größeren Speicherplatzbedarf, sondern führt auch zu einer Verlängerung der Erkennungszeit. Um diesen Aufwand zumindest teilweise zu reduzieren, werden vereinzelt Clusterverfahren angewendet. Dabei werden für mehrere sich ähnelnde Referenzmuster sogenannte Repräsentanten gebildet, die an die Stelle der von ihnen repräsentierten Referenzmuster treten [2].

Bei neueren Verfahren versucht man durch geeignete Transformationsverfahren, die sprechertypischen Anteile in den Merkmalsätzen zu unterdrücken, bzw. deren Einfluß zugunsten der wortbedeutungstragenden Bestandteile zu verringern [3].

3. Sprecheranpassung durch Sprechergruppenbildung

Insbesondere bei den zuerst genannten Verfahren im vorigen Abschnitt erfolgt die Berücksichtigung des Sprechereinflusses erst nach der abgeschlossenen Anlernphase (z.B. durch Clustern). Bei dem hier vorgestellten Verfahren wird der Sprechereinfluß bereits vor dem Beginn des eigentlichen Anlernvorganges bzw. vor der Erkennung berücksichtigt. Dadurch wird erreicht, daß sich die aufwendige Lernphase, dessen Aufwand um ein Vielfaches über dem Aufwand zur Erkennung liegt, vereinfacht. Der Grundgedanke dafür geht auf ein Verfahren zurück, das bei der automatischen Sprecheridentifikation bzw. Sprecherverifikation erfolgreich verwendet wurde. Bei diesem Teilgebiet der automatischen Sprachsignalverarbeitung werden seit längerer Zeit Langzeitfunktionen zur Erkennung herangezogen [4]. Dabei geht die Erkennung auf einen Ähnlichkeitsvergleich der Langzeitfunktionen unterschiedlicher Sprecher mit einem vorliegenden Muster (ebenfalls Langzeitfunktion) zurück. Durch die Bildung einer Rangfolge der ermittelten Ähnlichkeitsmaße zwischen vorliegendem Muster und aller in Frage kommenden Sprecher lassen sich die Ähnlichkeitsbeziehungen zum Muster ablesen. Hervorzuheben ist, daß die Langzeitfunktionen bei entsprechend langer Mittelungszeit unabhängig vom gesprochenen Text sind. Untersuchungen haben ergeben, daß 20 bis 30 s Sprachtext bereits eine stabile, den Sprecher beschreibende Langzeitfunktion liefert. Durch die Mittelung wird erreicht, daß die in jedem Sprachsignal enthaltene wortbedeutungstragende Information infolge des sich ständig ändernden Wortschatzes abgeschwächt wird. Dadurch, daß sich die ektosemantische Information (die den Sprecher beschreibende Information) über einen längeren Zeitraum kaum ändert, treten diese Merkmale durch die Mittelung stärker hervor. Daraus läßt sich schlußfolgern, daß Sprecher, die in der Rangfolge der Ähnlichkeitsmaße sehr weit vorn stehen, auch hinsichtlich der Artikulation und damit der Sprecherindividualität Ähnlichkeiten haben.

Bei dem hier beschriebenen Verfahren werden aufbauend auf diesen Ähnlichkeitsbeziehungen Sprechergruppen gebildet, für die im Anlernprozeß ein gemeinsamer Referenzdatensatz für die spätere Erkennung ermittelt wird. Ausgangspunkt für die Gruppenbildung ist die Ermittlung einer Ähnlichkeitsmatrix aller oder einer Vielzahl der für die Arbeit mit dem Erkennungssystem in Frage kommenden Nutzer. Als Ähnlichkeitsmaße eignen sich die in der Objekterkennung vielfach verwendeten Abstandsmaße, wie z.B. Euklidischer Abstand, City-Block-Distanz, Kreuzkorrelation, Mahalanobis-Distanz, Fisher-Kriterium u. dgl. [5]. Da frühere Untersuchungen gezeigt haben, daß diese Abstandsmaße bezüglich der hier geforderten Aussage nahezu

Tafel 1. Ähnlichkeitsmatrix von 10 männlichen Sprechern

Sprecher m \ Sprecher n	1	2	3	4	5	6	7	8	9	10
1	0,0	1,23	1,69	1,19	1,22	1,41	1,39	1,10	1,11	1,26
2	1,23	0,0	1,16	1,02	0,93	1,32	1,02	0,73	1,22	1,33
3	1,69	1,16	0,0	1,06	0,99	1,00	1,34	0,93	1,14	0,81
4	1,19	1,02	1,06	0,0	0,81	0,84	0,98	0,93	1,08	0,99
5	1,22	0,93	0,99	0,81	0,0	1,03	0,79	1,02	1,09	1,12
6	1,41	1,32	1,00	0,84	1,03	0,0	1,28	1,19	1,12	0,95
7	1,39	1,02	1,34	0,98	0,79	1,28	0,0	1,31	1,36	1,47
8	1,10	0,73	0,93	1,93	1,02	1,19	1,31	0,0	0,87	0,78
9	1,11	1,22	1,14	1,08	1,09	1,12	1,36	0,87	0,0	0,95
10	1,26	1,33	0,81	0,99	1,12	0,95	1,47	0,78	0,95	0,0

Tafel 2. Sprachschatz zur Bildung der Langzeitspektren

ACHT	EINS	KOMMA	NULL	START
BERECHNE	ERGEBNIS	COSINUS	PLUS	STOP
DREI	FÜNF	LOGARITHMUS	SAGE	VIER
DRUCKE	GETEILT	MAL	SECHS	WURZEL
DURCH	HOCH	MINUS	SIEBEN	ZU
E	KLAMMER	NEUN	SINUS	ZWEI

Tafel 3. Einzelerkennungsergebnisse (in %) als Ähnlichkeitsnachweis

angelernter Sprecher \ Systemnutzer	1	2	3	4	5	6	7	8	9	10
1	96,2	81,8	61,9	83,8	91,9	85,2	78,0	83,3	54,7	50,4
2	81,9	99,0	78,1	77,4	84,8	79,5	79,4	87,6	72,4	70,0
3	74,3	74,2	99,0	82,4	85,2	72,9	74,2	86,4	67,5	86,3
4	71,9	72,7	65,2	97,1	80,0	78,1	76,1	72,4	69,0	71,4
5	79,5	80,3	77,0	78,6	99,0	71,4	90,4	62,5	66,0	62,4
6	72,1	71,2	76,3	82,2	78,3	97,8	69,8	72,9	71,8	78,3
7	60,9	65,3	63,2	79,1	86,3	67,3	98,1	72,1	66,3	62,1
8	63,2	75,3	69,1	68,2	62,1	64,1	62,7	97,3	73,2	77,3
9	59,8	62,3	60,3	69,8	72,1	67,3	69,1	73,2	97,8	80,3
10	57,9	53,8	49,7	65,2	68,3	79,2	62,7	85,1	84,2	98,5

Tafel 4. Erkennungsergebnisse (in %) mit Referenzdaten von Sprecherpaaren

Sprecherpaar \ Systemnutzer	1	2	3	4	5	6	7	8	9	10
(2,8)	83,8	95,7	82,9	81,9	84,3	79,5	84,2	90,4	81,5	75,7
(5,7)	67,7	81,3	76,2	87,6	97,1	63,3	94,4	68,9	72,4	71,9

gleichwertig sind, wurde für den hier vorliegenden Fall die rechentechnisch sehr gut handhabbare City-Block-Distanz verwendet. Bei der City-Block-Distanz werden die Abstände der einzelnen Vektorkomponenten zweier Vektoren addiert:

$$CD_{xy} = \sum_{i=1}^{N} |x(i) - y(i)| . \qquad (1)$$

Dabei stellt $x(i)$ die i-te Komponente des Merkmalsvektors $\vec{x}$, $y(i)$ die i-te Komponente des Merkmalsvektors $\vec{y}$ und N die Anzahl der Vektorkomponenten dar. $\vec{x}$ und $\vec{y}$ sind die Langzeitvektoren der Sprecher X und Y. Als Langzeitvektoren wurden Langzeitmittelwertspektren verwendet. Sie lassen sich durch komponentenweise Mittelung der Kurzzeitspektralvektoren über einen längeren Zeitraum ermitteln:

$$x(i) = \frac{1}{M} \sum_{j=1}^{M} x_j(i) \quad \text{mit} \quad i = 1, 2, \ldots, N . \qquad (2)$$

$\vec{x}_j$ ist dabei eine Folge von Kurzzeitspektralvektoren über einen Zeitraum von $j = 1, \ldots, M$ Zeitfenster. Dabei bezeichnet i die

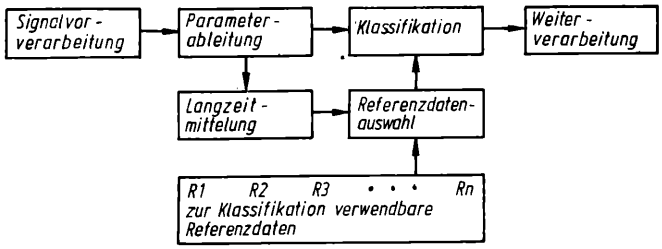


Bild 1. Sprecheradaptives Spracherkennungssystem

jeweils betrachtete Vektorkomponente. Die Folge von Kurzzeitspektralvektoren liefert eine 20kanalige Filterbank, die im Frequenzbereich von 70 bis 8600 Hz alle 12,5 ms einen Merkmalsvektor ermittelt. Die einzelnen Kanäle der Filterbank bestehen aus einem System von Bandpässen, die frequenzgruppenweise gestaffelt sind [6].

Tafel 1 zeigt die Ähnlichkeitsmaße, die von den Langzeitspektren von 10 männlichen Sprechern gebildet wurden. Als Datenmaterial wurde der in Tafel 2 angegebene Wortschatz verwendet. Beschränkt man sich bei der Sprechergruppenfindung zunächst auf Zweier-Gruppen, so hat nach Tafel 1 die Sprecherpaarung (2,8) die größte Ähnlichkeit (geringster Abstand). Dann folgen die Paarungen (8, 10) und (5, 7). Die in Tafel 3 angegebenen Einzelerkennungsergebnisse zeigen, daß die Sprecher, die nach Tafel 1 die größte Ähnlichkeit haben, auch hinsichtlich der erreichten Erkennungsraten die besten Ergebnisse liefern. Unter den vorliegenden Fremdsprechern liefert der Sprecher 8 das beste Ergebnis für ein System, das auf den Sprecher 2 angelernt wurde. Ein auf den Sprecher 5 angelerntes Erkennungssystem liefert für den Sprecher 7 das beste Fremderkennungsergebnis. Eine Verbesserung des Erkennungsergebnisses für den zweiten Sprecher erhält man, wenn man beim Anlernen beide Sprecher berücksichtigt. Dabei gehen Sprachrealisierungen beider in den Anlernprozeß ein. Tafel 4 zeigt die Erkennungsergebnisse für einen derartigen Fall. Ein Referenzdatensatz wurde mit Sprachrealisierungen der Sprecher 2 und 8 und der andere mit Realisierungen der Sprecher 5 und 7 gebildet. Untersuchungen haben gezeigt, daß Sprechergruppen mit 2 bis 3 Sprechern noch für die Spracherkennung akzeptable Ergebnisse liefern.

Bild 1 zeigt, wie das Verfahren der Sprechergruppenbildung in einem automatischen Spracherkennungssystem Berücksichtigung finden kann. Dem Erkennungssystem stehen mehrere Referenzdatensätze auf einem externen Speichermedium (Magnetplatte, Magnetband o. dgl.) zur Verfügung. Jeder Referenzdatensatz gilt für eine Sprechergruppe, die nach dem oben beschriebenen Verfahren gefunden wurde. Darüber hinaus sind neben diesen Referenzdaten die entsprechenden Langzeitfunktionen gespeichert. Bei der Arbeit mit dem Erkennungssystem wird parallel zum Erkennungsablauf die Langzeitfunktion des Nutzers ermittelt. Zur Erkennung wird dann der Referenzdatensatz herangezogen, dessen Langzeitfunktion die größte Ähnlichkeit mit der des Nutzers hat. Bei Sprecherwechsel wird erreicht, daß sich bereits nach etwa 10 Worten das Erkennungssystem an den neuen Sprecher angepaßt hat. Damit ist gesichert, daß nach Ablauf einer gewissen Adaptionszeit der für den jeweiligen Nutzer optimal zur Verfügung stehende Referenzdatensatz verwendet wird. In einer weiteren Ausbaustufe ist vorgesehen, daß durch Mittelwertbildung der zeitnormierten Sprachäußerung mit dem entsprechenden Referenzmuster eine zusätzliche Adaption erreicht wird, die insbesondere den bei jedem Sprecher typischen Veränderungen während des Sprechens Rechnung tragen.

Eingegangen am 15. September 1982 NaA 8991

Literatur

[1] *Levinson, S. E.; Rosenberg, A. E.; Flanagan, J. L.:* Evaluation of a Word Recognition System Using Syntax Analysis. Internat. Conference on Acoustics, Speech and Signal Processing 1977, S. 483–486

[2] *Rabiner, L. R.; Levinson, S. E.; Rosenberg, A. E.; Wilpon, J. G.:* Speaker Independent Recognition of Isolated Words Using Clustering Techniques. IEEE Trans. on Acoustics, Speech and Signal Processing. 27 (1979) 4, S. 336 bis 349

[3] *Matsumoto, H.; Wakita, H.:* Frequency Warping for Nonuniform Talker Normalization. Internat. Conference on Acoustics, Speech and Signal Processing 1979, S. 566–569

[4] *Markel, J. D.; Oshika, B. T.; Gray, A. H.* jr.: Long-Term Feature Averaging for Speaker Recognition. IEEE Trans. on Acoustics, Speech and Signal Processing 25 (1977) 4, S. 330–337

[5] *Affa, K.-H.:* Transformation und Merkmalbewertung. Diplomarbeit, TU Dresden, 1977

[6] *Affa, K.-H.:* Ein Verfahren zur Sprechererkennung und Sprechergruppenbildung. Dissertation, TU Dresden, 1982

Dipl.-Ing. *Karl-Heinz Affa*, Amt für Standardisierung, Meßwesen und Warenprüfung. 8027 Dresden, Georg-Schumann-Str. 7

Digitalfilterbank für Sprachanalyse auf einem KRS 4201

G. Weigert, Dresden

Mitteilung aus dem VEB Robotron Dresden, Zentrum für Forschung und Technik

Deskriptoren: Sprachanalyse, digitale Signalverarbeitung

1. Anforderungen und Zielstellungen

Grundlage für die automatische Erkennung der Sprache ist eine geeignete Verminderung des hohen Informationsflusses des Sprachsignals (Analyse) auf Werte, die dessen weitere Verarbeitung auf einem Mikrorechner in vertretbarer Zeit ermöglichen. Dafür werden bisher fast ausschließlich analoge Mittel eingesetzt. Gute Ergebnisse liefert eine Filterbank, die aus einer bestimmten Anzahl parallel arbeitender Filter besteht. Das Sprachsignal wird dabei in mehrere Frequenzintervalle zerlegt, deren Energiegehalt, integriert über Zeiträume von 10 bis 30 ms, als Merkmal zur Beschreibung der Sprachzeitfunktion verwendet werden kann.

Im Rahmen einer Diplomarbeit [1] wurde der Versuch unternommen, eine derartige Filterbank auf einem KRS 4201 zu programmieren. Dabei sollten vor allem Erfahrungen über Rechenzeitbedarf der Algorithmen und die erforderliche Koeffizientengenauigkeit gesammelt werden. Die Aufgabe wird dadurch erleichtert, daß die Anforderungen an die Filtergenauigkeit für Anwendungen in der automatischen Spracherkennung verglichen mit meßtechnischen Anwendungen nicht sehr hoch sind. Notwendig ist, das Frequenzband von 100 Hz bis 500 Hz in wenigstens 18 Kanäle zu zerlegen, was sich mit einer üblichen Terzfilterbank realisieren läßt. Die Filterflanken brauchen nicht steiler als 30 dB/Oktave zu sein, und geringe Abweichungen von der Mittenfrequenz bzw. Bandbreite idealer Terzfilter sind vertretbar ($\approx 5\%$).

Die Sprachzeitfunktion wurde im vorliegenden Beispiel mit 20 kHz bei einer Amplitudenauflösung von wahlweise 8 oder 12 bit abgetastet. Für jeden Abtastwert stand ein Maschinenwort (16 bit) des KRS 4201 zur Verfügung.

2. Digitale Filter

Unter einem digitalen Signalverarbeitungssystem soll im folgenden ein Algorithmus verstanden werden, der einer i.allg. unendlichen Eingangszahlenfolge $\langle x(k) \rangle$ gliedweise eine entsprechende Ausgangszahlenfolge $\langle y(k) \rangle$ zuordnet. Läßt sich der Algorithmus darüber hinaus durch eine lineare Differenzengleichung mit

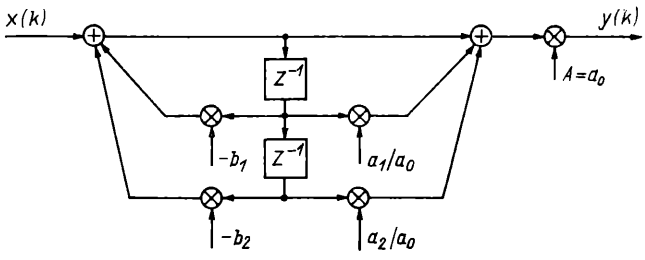


Bild 1. Digitalfilter zweiter Ordnung $H(z) = A \frac{1 + (a_1/a_0)\, z^{-1} + (a_2/a_0)\, z^{-2}}{1 + b_1\, z^{-1} + b_2\, z^{-2}}$

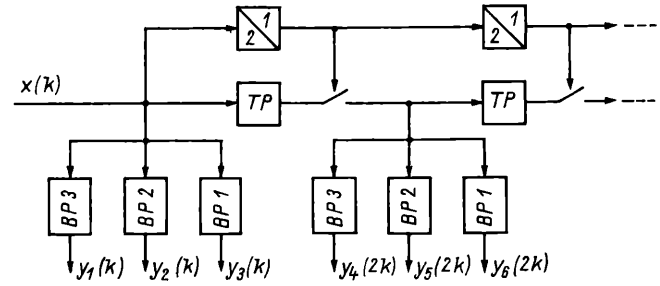


Bild 2. Blockschaltbild der digitalen Filterbank

konstanten Koeffizienten beschreiben, bezeichnet man das System auch als digitales Filter. Mit derartigen Algorithmen kann das Verhalten linearer analoger Systeme (z.B. Filter) simuliert werden, wenn die kontinuierliche Zeitfunktion durch Abtastung zu äquidistanten diskreten Zeitpunkten in eine Zahlenfolge umgewandelt wird. Die das analoge System beschreibende Differentialgleichung geht dann in eine Differenzengleichung über:

$$y(k) + \sum_{n=1}^{N} b_n\, y(k-n) = \sum_{n=0}^{N} a_n\, x(k-n)\,. \tag{1}$$

Wendet man auf Gl. (1) die Z-Transformation an, dann kann das digitale Filter auch durch seinen Übertragungsfaktor $H(z)$ beschrieben werden [3]:

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{n=0}^{N} a_n\, z^{-n}}{1 + \sum_{n=1}^{N} b_n\, z^{-n}}\,. \tag{2}$$

Im Bild 1 ist der Signalflußplan für ein digitales Filter 2. Ordnung dargestellt. Daraus geht hervor, daß der Algorithmus aus drei Operationen besteht: Addition, Multiplikation mit einer Konstanten, Umspeichern.

Die Arbeitsgeschwindigkeit eines digitalen Filters hängt wesentlich von den für diese Grundoperationen notwendigen Rechenzeiten ab. Die größten Schwierigkeiten bereitet dabei die Multiplikation, da Additions- und Umspeicheroperationen heute bereits von universellen Digitalrechnern ausreichend schnell ausgeführt werden. Der Multiplikationsoperation muß daher beim Entwurf digitaler Filter besondere Aufmerksamkeit geschenkt werden. Das gilt um so mehr für den KRS 4201, der über keine speziellen Multiplikationsbefehle verfügt.

Die im folgenden aufgezählten Besonderheiten digitaler Filter müssen beim Entwurf einer Filterbank außerdem berücksichtigt werden [4]:

1. Digitale Filter haben, bedingt durch die Abtastung der Signalfunktion, einen bezüglich der halben Abtastfrequenz f_A spiegelsymmetrischen Frequenzgang, der darüber hinaus mit der Periode f_A fortgesetzt erscheint. Von einer Äquivalenz zwischen analogem und digitalem System spricht man vereinbarungsgemäß dann, wenn die Frequenzgänge beider Systeme im Bereich $f = 0$ bis $f_A/2$ übereinstimmen. Das setzt allerdings voraus, daß die zu verarbeitenden Signale oberhalb $f_A/2$ bandbegrenzt werden.
2. Der Frequenzgang des digitalen Filters hängt nicht allein von den Koeffizienten seiner Differenzengleichung, sondern auch von der verwendeten Abtastfrequenz ab. Durch Halbierung der Abtastfrequenz wird die gesamte Frequenzachse um den Faktor 2 gestreckt.
3. Digitale Filter können im Multiplexbetrieb zur gleichzeitigen Verarbeitung mehrerer Zahlenfolgen eingesetzt werden. Voraussetzung dafür ist, daß vor Beginn der Verarbeitung einer neuen Zahlenfolge der Speicherinhalt des Filters gerettet wird.

3. Aufbau und Wirkungsweise der digitalen Filterbank

Bild 2 zeigt das Blockschaltbild der digitalen Filterbank. Die Bandpässe BP1, BP2 und BP3 sind Terzfilter, deren Grenzfrequenzen in folgender Beziehung zueinander stehen:

$$f_{oi} = \sqrt[3]{2}\, f_{ui} \quad \text{und} \quad f_{oi} = f_{ui+1} \quad \text{mit} \quad i = 1 \text{ bis } 3.$$

Auf diese Weise wird die gesamte Oktave f_{u1} bis $f_{o3} = 2 f_{u1}$ in drei Kanäle zerlegt. Die gleichen Filter lassen sich auch für die Analyse der nächstniedrigen Oktave $f_{u1}/2$ bis f_{u1} verwenden, wenn zuvor die Abtastfrequenz f_A geteilt wird. Das bewirken die 2:1-Teiler der Teilerkette, die jeweils einen Schalter betätigen, so daß nur jeder zweite Abtastwert durchgelassen wird. Für die gesamte Filterbank sind daher nur drei verschiedene Bandpässe notwendig, die im Multiplexbetrieb beliebig viele Oktaven verarbeiten können. Das Steuerprogramm der digitalen Filterbank muß dafür sorgen, daß beim Übergang zu einer neuen Oktave der Inhalt der Arbeitsspeicher der Filter nicht zerstört wird. Neben der Verminderung des Programmier- und Entwurfsaufwandes hat dieses Verfahren außerdem den Vorteil, Rechenzeit einzusparen, da sich bei der niedrigen Oktave gegenüber der höheren jeweils die Anzahl der zu verarbeitenden Abtastwerte je Zeiteinheit halbiert. Die Oktavtransformation der Filter durch wiederholte Halbierung der ursprünglichen Abtastfrequenz f_A birgt allerdings die Gefahr in sich, die durch die periodisch fortgesetzten Frequenzgänge digitaler Systeme bedingten Spiegelungsfehler zu erzeugen. Um das zu verhindern, ist vor jeder Halbierung der Abtastfrequenz eine erneute Bandbegrenzung der Abtastfolge auf die Hälfte der neuen Abtastfrequenz erforderlich. Das wird durch einen digitalen Tiefpaß TP bewirkt, der genau wie die Bandpässe im Multiplexbetrieb für beliebig viele Oktaven verwendet wird.

Die digitale Filterbank hat somit folgende Hauptbestandteile: 3 Terzfilter, Tiefpaß, Teilerkette, Steuerprogramm zur Organisation des Multiplexbetriebes.

Die auf dem KRS 4201 realisierte Filterbank besteht aus insgesamt 18 Kanälen, wodurch 6 Oktaven überstrichen werden. Am Ausgang eines jeden Kanals entsteht wieder eine Abtastfolge. Die drei Kanäle der höchsten Oktave geben zu jedem in die Filterbank eingegebenen Abtastwert je einen Abtastwert aus, die der nächstniedrigeren Oktave allerdings nur zu jedem zweiten Abtastwert usw. Alle Abtastwerte haben Festkommaformat. Die ausgegebenen Abtastfolgen repräsentieren die Ausgangszeitfunktionen äquivalenter analoger Filter. Für die Spektralanalyse ist es darüber hinaus erforderlich, in regelmäßigen Abständen einen Kurzzeiteffektivwert dieser Zeitfunktionen zu ermitteln. Es sei an dieser Stelle nur darauf hingewiesen, daß die Berechnung des Effektivwertes relativ zeitaufwendig ist. Rechentechnisch wesentlich einfacher ist dagegen eine Spitzenwertgleichrichtung [1] zu realisieren.

4. Entwurf der Filter

Das Hauptproblem beim Entwurf digitaler Filter besteht in der Realisierung der Multiplikationsoperation, da sie die notwendige Rechenzeit wesentlich bestimmt. Da die Multiplikation im Befehlsvorrat des KRS 4201 fehlt, wurden nur solche Filterkoeffizienten verwendet, die sich durch Binärzahlen mit möglichst wenig Ziffern (maximal 5) ausdrücken lassen. Die Multiplikation kann dann durch einige Verschiebe- und Additionsoperationen ersetzt werden.

Bandpässe

Grundelement der drei Bandpässe ist ein Potenzfilter zweiter Ordnung mit der Übertragungsfunktion

$$G(p) = \frac{\Delta\omega\, p}{p^2 + \Delta\omega\, p + \omega_m{}^2}\,. \tag{3}$$

Um die erforderliche Flankensteilheit zu erreichen, besteht jedes

Terzfilter aus der Kaskadenschaltung dreier gleichartiger Potenzfilter zweiter Ordnung. Man erhält den Übertragungsfaktor $H(z)$ des zugehörigen digitalen Filters durch folgende Substitution (bilineare Transformation [4]):

$$p = C \cdot \frac{1 - z^{-1}}{1 + z^{-1}}, \tag{4}$$

$$H(z) = A \cdot \frac{1 - z^{-2}}{1 + b_1 z^{-1} + b_2 z^{-2}}. \tag{5}$$

Der Frequenzgang des digitalen Filters wird allein durch die Konstanten b_1 und b_2 bestimmt. Es kann gezeigt werden, daß die Koeffizienten stabiler Systeme ausschließlich im schraffierten Dreieck der b_1, b_2-Ebene zu finden sind (Bild 3). Zwischen den Konstanten b_1 und b_2 sowie der absoluten Bandbreite Δf und der Mittenfrequenz f_m besteht der folgende elementare Zusammenhang:

$$\Delta f = \frac{C}{\pi} \cdot \frac{1 - b_2}{1 - b_1 + b_2}, \tag{6}$$

$$f_m = \frac{C}{2\pi} \sqrt{\frac{1 + b_1 + b_2}{1 - b_1 + b_2}}. \tag{7}$$

Die begrenzte Koeffizientenwortlänge (5 bit) erzeugt in der b_1, b_2-Ebene ein Raster, auf dessen Gitterpunkte die zulässigen Koeffizientenpaare (b_1, b_2) liegen. Ein als Entwurfshilfe gedachtes Programm ermittelt für alle zulässigen Koeffizientenpaare, die im Stabilitätsbereich liegen, die zugehörigen Werte für Δf und f_m und gibt diejenigen Paare aus, für die Bandbreite und Mittenfrequenz in einem zuvor festgelegten Toleranzbereich liegen. Dieses Programm berücksichtigt gleichzeitig die durch die Kaskadierung veränderte Bandbreite und berechnet auch den Verstärkungsfaktor A. Auf der Grundlage dieser Ergebnisse wurden die folgenden Filterkoeffizienten ermittelt:

BP1 $b_1 = 2^{-3} - 1$ $\quad b_2 = 2^{-1}(1 + 2^{-2})$ $\quad A = 0,18750$,
BP2 $b_1 = -2^{-1}$ $\quad b_2 = 2^{-1}(1 + 2^{-3})$ $\quad A = 0,21875$,
BP3 $b_1 = 0$ $\quad b_2 = 2^{-1}(1 - 2^{-4})$ $\quad A = 0,26564$.

Die zugehörigen Werte für Δf und f_m bei einer Abtastfrequenz von $f_A = 20$ kHz können dem Bild 4 entnommen werden.

Tiefpaß

Grundelement des Tiefpasses ist ebenfalls ein Potenzfilter. Allerdings konnte die notwendige Flankensteilheit, die wesentlich durch das Verhältnis der Mittenfrequenz des obersten Terzfilters zur Abtastfrequenz und durch die angestrebte Amplitudenauflösung von 12 bit bestimmt wird, nur durch die Hintereinanderschaltung dreier gleichartiger Potenzfilter vierter Ordnung erreicht werden. Da sich ein Potenzfilter vierter Ordnung in zwei Filter zweiter Ordnung zerlegen läßt, kann der Entwurfsprozeß in ähnlicher Weise wie zuvor beschrieben erfolgen [2]. Den Frequenzgang des Tiefpasses zeigt Bild 4.

5. Zusammenfassung

Die Funktion der digitalen Filterbank konnte mit realen Sprachzeitfunktionen nachgewiesen werden (Amplitudenauflösung 12 bit, Abtastfrequenz 20 kHz). Die Koeffizientengenauigkeit (5 bit) erwies sich in diesem Fall als ausreichend. Für die Filterung eines 12,8 ms langen Sprachsignalabschnittes (256 Abtastwerte) benötigt der KRS 4201 etwa 1 s. Damit ist eine Echtzeitverarbeitung nicht möglich.

Seit einiger Zeit werden auf dem Weltmarkt allerdings erste programmierbare Signalprozessorschaltkreise [5] [6] angeboten. Sie zeichnen sich vor allem durch die Fähigkeit aus, Multiplikationen in wenigen hundert Nanosekunden ausführen zu können. Dadurch wird es möglich, eine digitale Filterbank mit vergleichbaren Parametern auf einem Chip zu realisieren, die Sprachzeitfunktionen in Echtzeit verarbeitet.

Eingegangen am 15. September 1982 NaA 8992

Bild 3. b_1, b_2-Ebene mit Stabilitätsbereich (schraffiert)

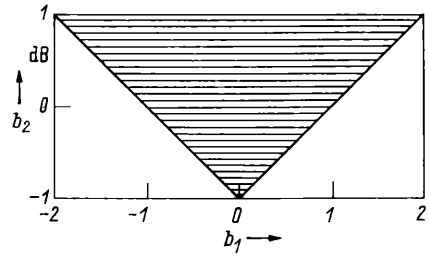
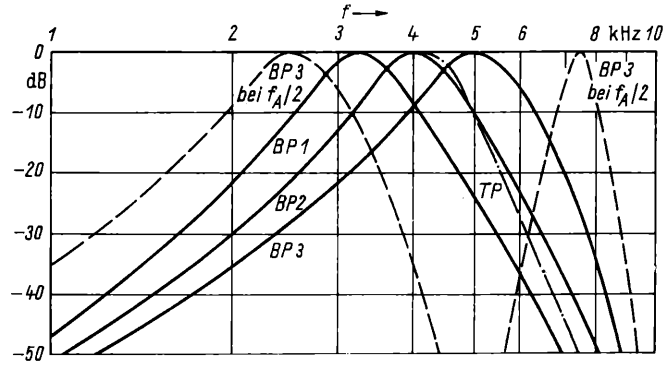


Bild 4. Frequenzgangkurven bei $f_A = 20$ kHz



Literatur

[1] *Henker, R.:* Digitale Filterbank. Diplomarbeit. TU Dresden, 1981
[2] *Stanley, W. D.:* Digital Signal Processing. Reston: A Prentics-Hall Company 1975
[3] *Weigert, G.:* Beschreibung digitaler Signalverarbeitungssysteme. Preprint 09-09-81, TU Dresden
[4] *Weigert, G.:* Entwurf digitaler Filter. Preprint 09-10-81, TU Dresden
[5] *Capace, R. P.:* Digital n-MOS chip handles analog signals at record 200 kHz. Electronics 53 (1980) 26, S. 79–80
[6] *Posa, J. G.:* Signalprocessor spur changes. Electronics 53 (1980) 4, S. 100

Dipl.-Ing. *Gerald Weigert*, VEB Robotron ZFT, 8012 Dresden, Postschließfach 330

Multiples Modell zur Beschreibung subjektiver Klangattribute

K. Haller, KDT, Oelsnitz; F. Blutner, Zwota

Mitteilung aus dem VEB Kombinat Musikinstrumente Markneukirchen/Klingenthal

Deskriptoren: Schall, Schallmessung, Klang, akustische Eigenschaften, akustische Meßverfahren, Qualitätsbewertung, Modell

1. Lösungsansatz

Die Beschreibung des funktionalen Zusammenhangs zwischen den physikalischen Schallparametern und den subjektiven Wahrnehmungsgrößen ist ein Kernproblem der musikalischen Akustik.

Auf der Grundlage einer Analyse/Synthese-Konzeption wurde von dem in Bild 1 gezeigten Lösungsansatz ausgegangen. Dabei sind zunächst im Hörtest Messungen am Modelloriginal vorzunehmen. Resultat ist die analytische Beschreibung der angebotenen Klänge in Form einer Reihe ausgewählter Klangattribute. Diese Attribute sind die Zielgrößen $ZG_{\text{subj.}}$ der statistischen Optimierung.

Zur Approximation dieser Zielgrößen aus den physikalischen Signaleigenschaften wurde das im Bild oben dargestellte dreistufige Meßsystem eingesetzt:

1. Analysator (Terzfilter, Frequenzgruppenfilter, Oktavfilter, Diskriminationsfilter, Fast-Fourier-Analysator, Linear-Prediction-Analysator).
2. Eigenschaftsselektor (Tiefpaß-Integrator, Envelopemeter [1], Tonhöhendetektor [2], Transientdetektor [3]).
3. Diskriminator (verschiedene Varianten des linearen Optimierungsmodells).

Der Diskriminator ermittelt in einer Lernphase eine optimale Verknüpfung zwischen den eingangsseitigen objektiven Einflußgrößen E_i und den vorgegebenen Zielgrößen.

Im Ergebnis der Lernphase liefert der Diskriminator in einer nachfolgenden Kannphase objektive Zielgrößen. Diese resultierenden Meßgrößen $ZG_{\text{obj.}}$ approximieren die subjektiven Zielgrößen $ZG_{\text{subj.}}$ mit einem minimalen Restfehler.

2. Multiples Optimierungsmodell

Bei der mathematischen Modellierung des Diskriminators wird von folgendem linearem Modellansatz ausgegangen:

$$\sum_{i=1}^{m} a_i E_{ij} = ZG_{\text{obj. j}} \quad j = 1 \text{ bis } N$$

$$\sum_{j=1}^{N} (ZG_{\text{subj. j}} - ZG_{\text{obj. j}})^2 \rightarrow \min.$$

Dieses Modell ermittelt im Ergebnis der optimalen Anpassung objektiver und subjektiver Zielgrößen die Gewichtskoeffizienten a_i. Auf Grund des häufig auftretenden korrelativen Zusammenhangs zwischen den Einflußgrößen E_i (Multikollinearität), liefert die Optimierung trotz gleichbleibend guter Approximation der subjektiven Zielgröße für verschiedene Stichproben aus der gleichen Grundgesamtheit unterschiedliche Gewichte. Diese Sensibilität, die sich in der Hypothesenprüfung durch große Konfidenzintervalle für die a_i zeigt, weist auf verdeckte Freiheitsgrade im Ergebnis hin, deren Fixierung nur durch Einbringung zusätzlichen Wissens möglich ist. Zur Auffindung dieser Freiheitsgrade bestehen folgende Möglichkeiten:

1. Unmittelbare Darstellung der Wechselwirkung zwischen den Gewichtskoeffizienten (grafische Darstellung der Beziehungen zwischen den a_i bei Variation ausgewählter Gewichte),

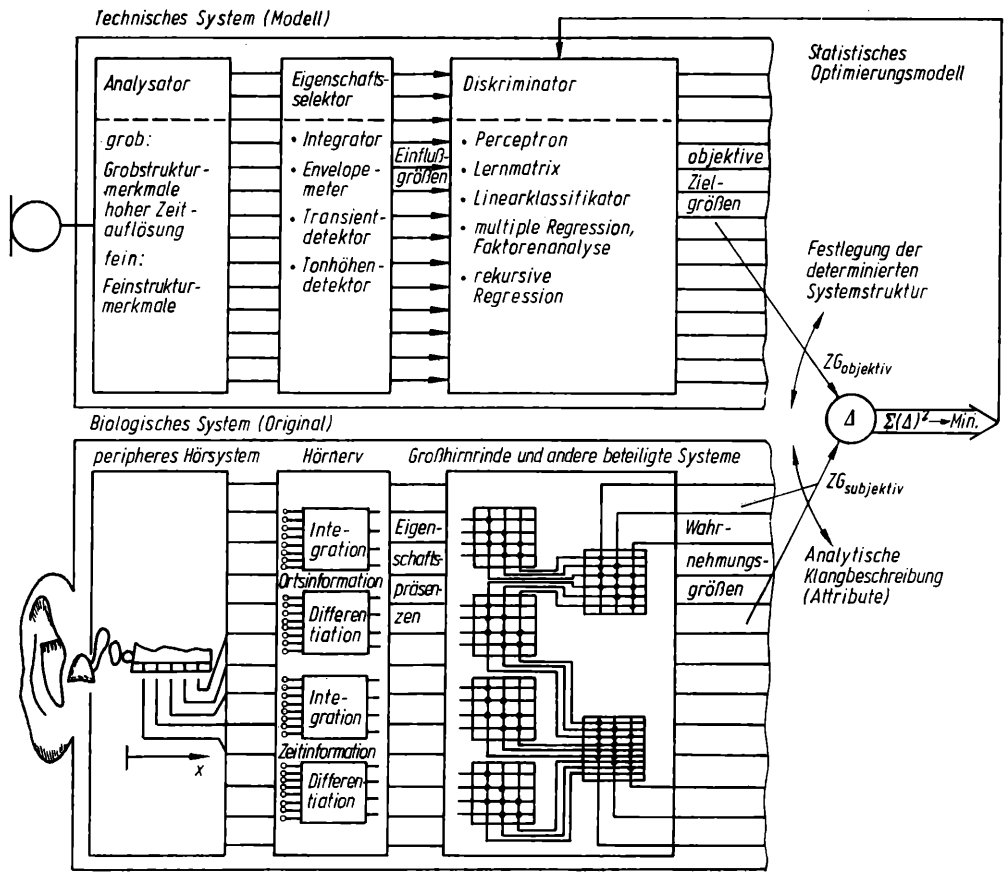


Bild 1. Kybernetischer Lösungssatz zur hörgerechten Erfassung musikalischer Töne

Tafel 1. Experimentelle Randbedingungen ausgewählter Beispiele der Qualitätsbeurteilung von Musikinstrumenten

Charakterisierung der untersuchten Klänge		Charakterisierung der subjektiven Zielgrößen $ZG_{subj.}$		Charakterisierung der berechneten Zielgrößen $ZG_{obj.}$		Korrelation $ZG_{subj.} - ZG_{obj.}$
Instrument	Tonbereich	Attribut	Größe der Versuchsgruppe	Eigenschaftsselektor	dynamische Wichtung spektraler Komponenten	
Akkordeon	Baß $F - d^1$	Lautstärke	24 VP_n*)	Lautstärkeanalysator 8051 A	frequenzunabhängig	Bild 2a)
				Terz-Echtzeitanalysator 01012	mehrgipflige Gewichtsfunktion	
	Mittellage $a - a^2$	Schärfe	24 VP_n	Schärfe nach *Bismarck*	monoton steigende Gewichtsfunktion	Bild 2b)
	Diskant $d^2 - d^4$			Terz-Echtzeitanalysator 01012	mehrgipflige Gewichtsfunktion	
				Linear-Prediction-Analyse	mehrgipflige Gewichtsfunktion	
Gitarren	$A - e^1$	Lautstärke	16 VP_n	A-bewerteter Schalldruckpegel $dB_{A,F}$	ohne	Bild 3a)
				Lautstärkeanalysator 8051 A	frequenzunabhängig	
				Terz-Echtzeitanalysator 01012 Maximalwertspektrum	mehrgipflige Gewichtsfunktion	
				Envelopemeter	mehrgipflig, rekursive Optimierung	
		Schärfe	16 VP_n	Obertonhaltigkeit nach *Lieber*	ohne	Bild 3b)
				Schärfe nach *Bismarck*	monoton steigende Gewichtsfunktion	
				Terz-Echtzeitanalysator 01012 Maximalwertspektrum	mehrgipflige Gewichtsfunktion	
				Envelopemeter	mehrgipflige rekursive Optimierung	
Piano	$A_1 - Q^4$	subjektive Klangdauer	16 VP_n	Abklingzeit des Gesamtsignals	ohne	Bild 3c)
				Abklingzeiten spektraler Komponenten	mehrgipflige Gewichtsfunktion	
				Terz-Echtzeitanalysator 01012 Maximalwertspektrum	mehrgipflige Gewichtsfunktion	
			*) VP_n Versuchspersonen	Pegelklassiereinrichtung	mehrgipflige Gewichtsfunktion	

2. Transformation der Einflußgrößen nach Parametern der Stichprobe (Hauptachsentransformation),
3. Funktionaltransformation der Einflußgrößen auf der Grundlage geeigneter Verfahren der Signalanalyse (z.B. LPC-Transformation).

Hiervon ausgehend können die aufgedeckten Freiheitsgrade durch folgende Zusatzkenntnisse eingeschränkt werden:

1. Physikalisches Wissen über die Schallerzeugung (Sprecher, Musikinstrument),
2. Kenntnisse über die Einflußgrößen (z.B. aus der Signalanalyse bekannte Zusammenhänge zwischen den Einflußgrößen),
3. Psychoakustische Kenntnisse über die Schallperzeption (Hörer).

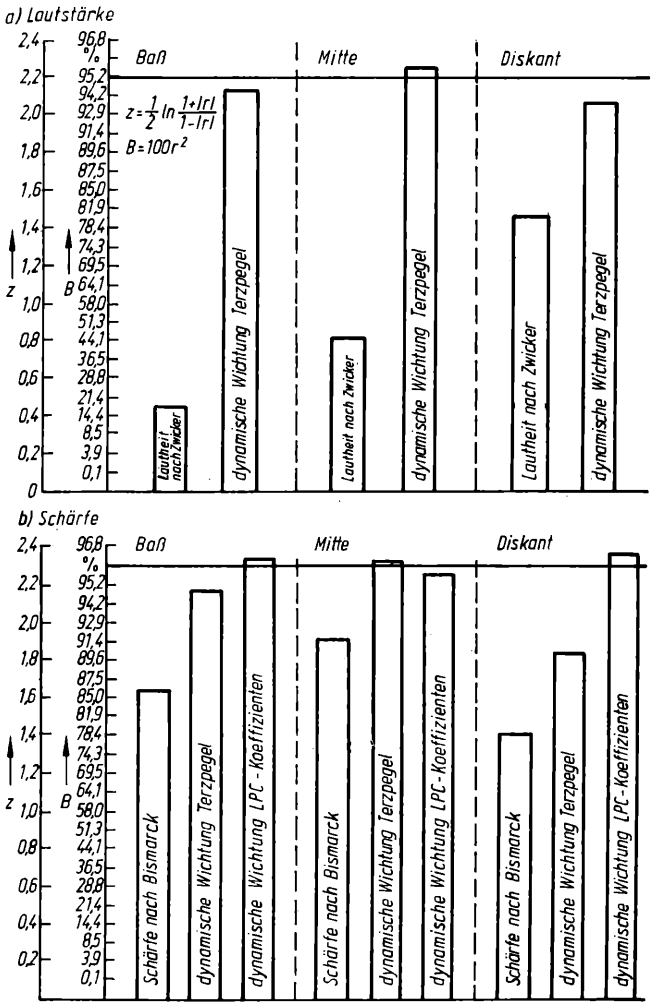
3. Anwendungsbeispiele

An Hand von Beispielen werden einige praktische Ergebnisse erläutert. Tafel 1 gibt zunächst einen Überblick zu den experimentellen Randbedingungen. Als Reize fungieren musikalische Töne. Dabei wurden die Akkordeonklänge über eine maschinelle Anregung erzeugt, während die Gitarren- und Pianoklänge von einem professionellen Interpreten angespielt wurden.
Diese Klänge wurden im Hörtest jeweils paarweise angeboten (alle möglichen Kombinationen der Testinstrumente bei konstant gehaltener Tonhöhe, subjektive Skalierung in 5 Stufen). Hierzu ist zu bemerken, daß die Unterschiede zwischen den Testinstrumenten nur gering waren (die Instrumente entstammen jeweils derselben Gebrauchswertklasse). Die Situation war also etwa die, vor der ein Käufer steht, wenn er die Wahl zwischen etwa gleich guten Instrumenten hat.
Die subjektiven Zielgrößen wurden mit verschiedenen objektiv gemessenen oder berechneten Zielgrößen verglichen. Die teilweise recht umfangreichen und komplizierten Rechenvorschriften können hier nicht im Detail wiedergegeben werden.
Als erstes Beispiel seien Untersuchungen zu dem wohl etablierten Attribut „Lautstärke" an quasistationären Akkordeonklängen erläutert. Interessant war hierbei zunächst die Frage, wie genau das Lautheitsmeßverfahren nach *Zwicker* [4] die unter den Bedingungen des Paarvergleichs ermittelten subjektiven Skalenwerte approximiert. Dabei ist zu beachten, daß Akkordeonklänge bereits eine recht komplizierte Spektralstruktur mit teilweise deutlich ausgeprägten Maxima und Minima haben. Es zeigte sich, daß unter diesen Bedingungen z.T. nur unbefriedigende Korrelationen zu den subjektiven Zielgrößen erreicht werden (Bild 2a). Im Bild wurde die Bestimmtheit $B = 100\,r^2$ (r Korrelationskoeffizient) in der z-transformierten Darstellung aufgetragen. So werden z.B. in der Baßlage nur etwa 15% der Variabilität der subjektiven Lautstärkeskala über die Lautheit nach *Zwicker* richtig beschrieben. In der Mittellage wurde die Bestimmtheit von 44% und in der hohen Tonlage $\approx 81\%$ ermittelt. Der letzte Wert entspricht einem Korrelationskoeffizienten von 0,9.

Während nach dem psychoakustischen Lautheitsmodell die Beziehung zwischen Erregung E und spezifischer Lautheit N' nach einem Potenzgesetz mit einem Exponenten k unabhängig von der Frequenz beschrieben wird, wurden stark frequenzabhängige Gewichtsfunktionen ermittelt. Im genannten Beispiel ergaben sich besonders bei tiefen Tönen deutlich mehrgipflige Gewichtsfunktionen, während in den hohen Tonlagen mehr glatte Verläufe auftraten.

Die Berücksichtigung dieser Zusammenhänge im Software-Simulationsmodell ergab ausgezeichnete Approximationsgüten der subjektiven Lautstärkeskala (Bild 2a). Im allgemeinen entsprechen diese der Reproduzierbarkeit der aus 25 Versuchspersonen bestehenden Hörergruppen von $B = 95\%$ (ausgezogene Gerade). Das zweite Beispiel zeigt für die Klangschärfe entsprechende Ergebnisse (Bild 2b). Während glatte Gewichtsfunktionen (z.B. monoton steigende Wichtung $g(z)$ nach *Bismarck* [5]) Bestimmtheiten um 80% ergaben, bringen hiervon mehr oder weniger „s-förmig" abweichende Gewichtsfunktionen Bestimmtheiten um 90%. Weitere Verbesserungen bis auf 96% sind dann nur noch durch Berücksichtigung tonaler Merkmale möglich. Als geeignetes Verfahren hat sich dabei die aus der Sprachforschung bekannte Linear-Predictions-Methode [3] erwiesen.

An einem nächsten Beispiel seien einige Untersuchungen zur Lautstärke typisch nichtstationärer musikalischer Töne erläutert (Bild 3a). Zunächst zeigt sich auch hier für Gitarren und Pianos der Effekt mehrgipfliger spektraler Gewichtsfunktionen.



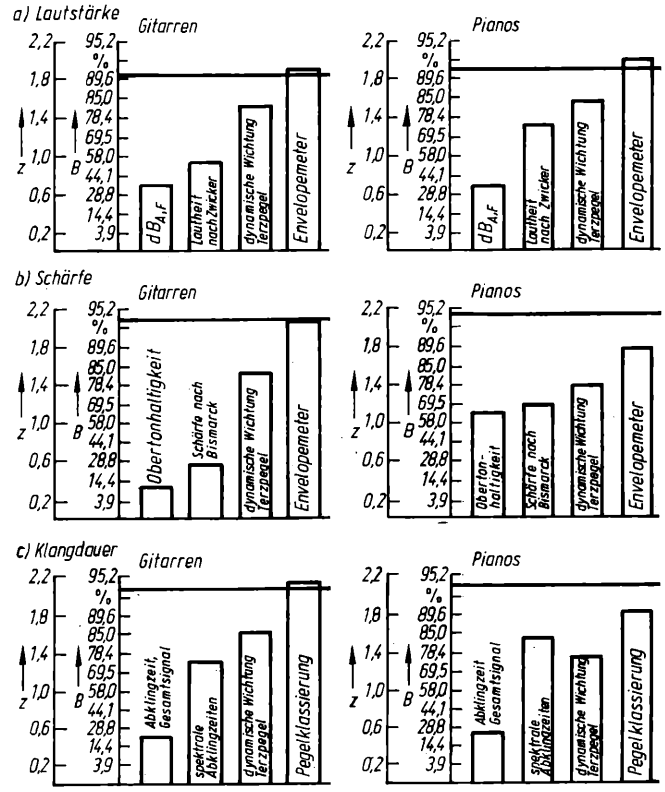


Bild 3. Beispiele zur Approximation subjektiver Skalen bei der Qualitätsbeurteilung instationärer musikalischer Töne

Bild 2. Beispiele zur Approximation subjektiver Skalen bei der Qualitätsbeurteilung von Akkordeonklängen

Die erhaltenen Approximationsgüten von 83 bis 84% waren jedoch nicht befriedigend. Deshalb wurde ein Spezialgerät zur Erfassung der schnellen Einschwingvorgänge entwickelt. Bei Einsatz dieses Envelopemeters ergaben sich deutlich bessere Bestimmtheiten zwischen 91 und 93%. Das bedeutet, daß die Lautstärke komplexer instationärer Klänge nicht nur eine spektrale Intensitätsgröße ist, sondern auch von der Art des Klangeinsatzes (hart oder weich) beeinflußt wird.

Wie in Bild 3b) gezeigt, wird auch die Klangschärfe von den Einschwingvorgängen beeinflußt, wie überhaupt bei allen noch untersuchten Attributen (Klarheit, Brillanz, Härte usw.) ein starker Einfluß dieser Vorgänge festzustellen war.

Bild 3c) zeigt Ergebnisse zur Klangdauerskala. Hier bringen Kennwerte des Gesamtsignals (Pegel-Zeitverlauf) nur ungenügende Korrelationen. Erst die Erfassung spektraler Kennwerte wie „Abklingzeit der Teiltöne" aber auch „Maximalwertspektrum" kann als Indikator für die subjektive Dauer benutzt werden. Die besten Ergebnisse liefert jedoch ein Kennwert, der durch Anwendung des Verfahrens der Pegelklassierung gewonnen wurde.

4. Zusammenfassung

Es wurde ein mathematisches Modell zur Quantifizierung des komplexen Zusammenhangs zwischen physikalischen Schallparametern und subjektiven Klangattributen vorgestellt.

Dabei konnte durch Berücksichtigung psychoakustischer Erkenntnisse im determinierten Strukturmodell eine aufwandsgünstige Lösung erzielt werden. Unter Nutzung multipler statistischer Optimierungsverfahren ist in diesem Fall eine sehr genaue Approximation der im Hörtest ermittelten subjektiven Attribute möglich. Der Approximationsfehler liegt in der Größenordnung der bei Hörversuchen auf Grund von Lern- und Adaptionseinflüssen stets auftretenden systematischen Fehler.

Eingegangen am 15. September 1982 NaA 8993

Literatur

[1] *Blutner, F.:* Verfahren zur Erfassung der Zeitenvelope von Musik- und Sprachklängen. DD-WP 210353. Int. Cl. G 10 L 1/02

[2] *Tscheschner, W.; Pohlink, M.:* Schaltungsanordnung zur Ermittlung der Stimmband-Grundfrequenz. DD-WP 49353. Int. Cl. G. 10k 1/01

[3] *Markel, J. D.; Gray, A. H.:* Linear prediction of speech. Berlin, New York: Springer-Verlag 1976

[4] *Zwicker, E.:* Procedure for calculating loudness of temporally variable sounds. J. Acoust. Soc. Am. 62 (1977) S. 675–682

[5] *Bismarck, G. v.:* Extraktion und Messung von Merkmalen der Klangfarbenwahrnehmung stationärer Schalle. Dissertation TU München, 1972

Dipl.-Math. *Konrad Haller* Kombinat Musikinstrumente, 9900 Plauen. Engelstr. 8
Dr.-Ing. *Friedrich Blutner*, Institut für Musikinstrumentenbau, 9657 Zwota, (Vogtland) Klingenthaler Str. 42

Sprachdaten-Gewinnung nach dem Linear-Predictive-Coding-Verfahren

G. Jaeger, KDT, Dresden

Mitteilung aus dem VEB Robotron, Zentrum für Forschung und Technik, Dresden

Deskriptoren: Sprachanalyse, Sprachsynthese; Sprachdatenverdichtung; Informationsfluß; Anregungssignale, Übertragungsfunktionen, Sprachsignal, Grundfrequenz; rechentechnische Realisierung

Die Bedeutung der LPC (Linear-Predictive-Coding) für die Sprachsynthese liegt in seiner datenkomprimierenden oder redundanzvermindernden Eigenschaft. Dadurch ist es möglich, PCM-codierte Sprache so stark zu komprimieren, daß der Informationsfluß von 80 kbit/s auf 1 bis 2 kbit/s reduziert wird, ohne daß eine wesentliche Einbuße an Sprachqualität bei der wiedergewonnenen synthetischen Sprache zu verzeichnen ist.
Die datenkomprimierende Eigenschaft hat ihre Ursache darin, daß die im Sprachsignal enthaltene Information aufgespalten wird in den Anteil, der das Übertragungsverhalten des Artikulationstraktes beschreibt und in das Anregungssignal. Während das Anregungssignal nur die Merkmale Amplitude und Grundfrequenz enthält, repräsentiert das Übertragungsverhalten des Artikulationstraktes die eigentliche sprachrelevante Information. Sie wird in Form von LPC-Koeffizienten gewonnen. Die Datenreduktion resultiert speziell daraus, daß sich das Übertragungsverhalten entsprechend der Lautbildung der Sprache nur relativ langsam ändert. Dementsprechend erfolgt die Berechnung der LPC-Koeffizienten für Zeitabschnitte, innerhalb derer die Sprachzeitfunktion als quasistationär zu betrachten ist, was praktisch einer Abtastung des Übertragungsverhaltens entspricht.
Die theoretische Begründung des LPC-Verfahrens erfolgte vor etwa 10 Jahren durch *Atal, Makhoul, Markel* u.a. [1] [2] [3] [4]. Seither hat das Verfahren international eine sehr starke Verbreitung und Anwendung gefunden, was sich speziell im Angebot einer Vielzahl von Sprachsyntheseschaltkreisen äußert. Für die LPC-Analyse, d.h. die Gewinnung der LPC-Daten zur Ansteuerung der Syntheseschaltkreise, ist bisher keine Geräte- oder Schaltkreislösung im Angebot. Die bisherige Verfahrensweise ist vielmehr so, daß die Hersteller von Sprachsynthesegeräten bereits bestimmte Wortschätze für verschiedene Anwendungen fest generieren. Das hat seine Ursache offenbar darin, daß die LPC-Analyse ein relativ komplizierter Prozeß ist, der sich infolge der Zufälligkeiten im natürlichen Sprachsignal nicht vollständig automatisieren läßt.
Die folgenden Ausführungen beschreiben die Arbeitsweise eines Programmsystems für die LPC-Datengewinnung auf der Basis des Wirtsrechners KRS 4201.

1. Sprachdateneingabe

Ausgangspunkt für die Sprachverarbeitung nach dem LPC-Verfahren ist die Digitalisierung der Sprachzeitfunktion, die zunächst auf Analogmagnetband gespeichert vorliegt. Dazu steht für die vorliegenden Untersuchungen der Analysemeßplatz AMP1 als Labormuster zur Verfügung. Die AD-Wandlung erfolgt mit 12 bit Quantisierungsbreite und wahlweise 8 oder 10 kHz Abtastfrequenz. Im Zusammenwirken mit dem Dateneingabeprogramm DAT 4200 werden die Sprachdaten in den Hauptspeicher des KRS 4201 eingelesen. Die Gesamtdatenmenge entspricht einer Sprachprobe von 2,5 s Länge. Nach der Speicherung wird die Sprachprobe kontrollgelesen und fortlaufend zyklisch wieder ausgegeben. Dabei ermöglicht der Meßplatz AMP1 mit dem Programm DAT 4200 eine genaue Begrenzung der Sprachprobe auf den interessierenden und anschließend zu verarbeitenden Datenbereich.

2. LPC-Programmsystem

Das LPC-Programmsystem besteht aus den Teilprogrammen LPCA (Analyse) und LPCS (Synthese). Nach der Dateneingabe berechnet das Analyseprogramm LPCA aus der Sprachzeitfunktion die LPC-Koeffizienten. Zusätzlich werden in diesem Programm als notwendige Informationen für die Synthese der Effektivwert und die Grundfrequenz mitberechnet. Das Syntheseprogramm LPCS hat die Aufgabe, aus den LPC-Koeffizienten wieder die Sprachzeitfunktion zu berechnen. Indem sie anschließend über den Analysemeßplatz AMP1 digital/analog rückgewandelt wird, ist zugleich eine akustische Qualitätskontrolle möglich.
Sowohl die Dateneingabe als auch die Ausgabe der synthetisierten Sprachdaten erfolgen in Echtzeit. Das Analyseprogramm benötigt die 300fache und das Syntheseprogramm die 50fache Echtzeit.

2.1. Analyseprogramm LPCA

2.1.1. LPC-Koeffizientenberechnung

Die Analyse des digitalisierten Sprachsignals erfolgt zeitfensterweise mit einer Breite und einem Zeitfensterabstand von 20 ms. Innerhalb dieser Zeit läßt sich die Sprachzeitfunktion als quasistationär betrachten. Vor der Abarbeitung des LPC-Algorithmus werden zunächst folgende Vorverarbeitungsstufen durchlaufen:

— Effektivwertberechnung
— Amplitudennormalisierung
— Preemphasierung des Sprachsignals
— Hamming-Bewertung.

Die Amplitudennormalisierung bewirkt eine Genauigkeitsverbesserung der LPC-Berechnung bei Sprachsegmenten mit geringem Effektivwert. Durch die Preemphasierung wird der im Sprachsignal für stimmhafte Laute enthaltene $1/\omega$-Frequenzgang eliminiert, da er keine Eigenschaft des Artikulationstraktes, sondern des Anregungssignals ist. Die Hamming-Bewertung hat die Aufgabe, die durch das Zeitfenster bedingte Spektralverfälschung zu minimisieren.
Der Grundalgorithmus für die LPC-Analyse ist der von *Markel/Gray* [4] angegebene Rekursionsalgorithmus mit den Autokorrelationskoeffizienten $r(i)$ des Sprachsignals $s(n)$ als Eingangsgröße. Die Autokorrelationskoeffizienten $r(i)$ werden aus N Stützwerten des betrachteten Zeitfensters nach folgender Gleichung berechnet:

$$r(i) = \sum_{n=0}^{N-i-1} s(n)\, s(n+i) \quad \text{mit} \quad i = 0 \text{ bis } M. \tag{1}$$

M ist dabei die Anzahl der zu berechnenden LPC-Koeffizienten, die im vorliegenden Fall mit $M = 10$ festgelegt ist, womit das Sprachsignal hinreichend genau beschrieben ist.
Der Rekursionsalgorithmus liefert die Reflexionskoeffizienten k_m und die Prediktorkoeffizienten a_{mi} für $m = 0$ bis M nach folgenden Gleichungen:

$$k_{m+1} = -\frac{1}{\alpha_m} \sum_{i=0}^{m} r(m+1-i)\, a_{mi}; \quad k_1 = -\frac{r(1)}{r(0)} \tag{2}$$

$$\alpha_{m+1} = \alpha_m - k_{m+1}^2 \cdot \alpha_m; \qquad \alpha_0 = r(0) \tag{3}$$

$$\left.\begin{aligned} a_{m+1,0} &= 1 \\ a_{m+1,i} &= a_{mi} + k_{m+1} \cdot a_{m,(m+1-i)}; \qquad i = 1, \ldots, m \\ a_{m+1,m+1} &= k_{m+1} \end{aligned}\right\}. \tag{4}$$

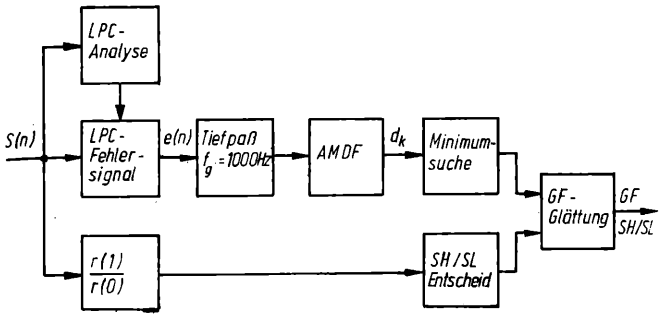


Bild 1. Blockschaltbild des Grundfrequenzalgorithmus

Jeweils ein Satz von M Reflexionskoeffizienten k_m oder Prediktorkoeffizienten a_m repräsentiert eine Übertragungsfunktion des Artikulationstraktes für das jeweilige Sprachsegment.
Die Berechnung der LPC-Koeffizienten erfolgt mit Gleitkommagenauigkeit. Für die Synthese ist jedoch eine Genauigkeit von 10 bit ausreichend. Damit ergibt sich zusammen mit den Zusatzinformationen Effektivwert und Grundfrequenz eine Informationsdichte von etwa 6 kbit/s.
Eine weitere Datenverdichtung wird erreicht, indem die Koeffizienten mit Hilfe einer Literaltabelle codiert werden. Dabei ist zu beachten, daß die Koeffizienten nicht gleichverteilt sind und daß ihre Variationsbreite unterschiedlich ist, woraus sich eine unterschiedliche Anzahl von Literalen je Koeffizient ergibt. Durch eine optimale Codierung läßt sich die Informationsdichte auf 2,5 kbit/s reduzieren. Weitere Einsparungen sind im Mittel über längere Sprachproben möglich bei Berücksichtigung von Sprechpausen, längeren stationären Phasen und stimmlosen Lauten [5].

2.1.2. Grundfrequenzbestimmung

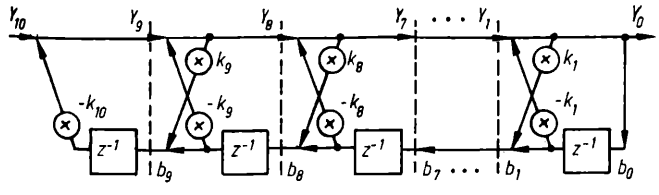
Die Bestimmung der Grundfrequenz im Analyseprozeß ist notwendig, um einerseits zwischen stimmhaften und stimmlosen Lauten zu unterscheiden, da bei der Synthese die Anregung entsprechend umgeschaltet werden muß. Andererseits stellt die Grundfrequenz ein prosodisches Merkmal dar, das für die Natürlichkeit der Sprache von großer Bedeutung ist.
Als Grundfrequenzalgorithmus wird das von *Un/Yang* [6] vorgeschlagene Verfahren der gemittelten Differenzfunktion (AMDF) angewandt.
Im Bild 1 ist das Blockschaltbild für den Grundfrequenzalgorithmus angegeben. Für die Bestimmung der Grundfrequenz wird ein Zeitfenster von 40 ms verwendet. Der Fensterfortschritt ist wie bei der Analyse 20 ms. Als Ausgangssignal für die Grundfrequenzbestimmung wird das LPC-Fehlersignal benutzt, das als Anregungsfunktion des Sprachsignals zu betrachten ist. Das Fehlersignal $e\,(n)$ ergibt sich mit dem Direktformfilter und den Prediktorkoeffizienten a_i aus der Gleichung

$$e\,(n) = \sum_{i=0}^{M} a_i\, s\,(n-i); \qquad \begin{matrix} a_0 = 1 \\ M = 10 \end{matrix} \qquad (5)$$

Zur Tiefpaßfilterung dient ein digitales Filter mit einer exakten Nullstelle bei 1000 Hz.
Die gemittelte Differenzfunktion d_k ist eine äquivalente Form zur Autokorrelationsfunktion, wobei statt des Produktes die Differenz zweier Stützwerte gebildet wird. Die Gleichung lautet

$$d_k = \frac{\sum_{i=1}^{N} |e_n - e_{n-k}|}{\sum_{i=1}^{N} |e_n|}; \qquad k = 0, \ldots, N. \qquad (6)$$

Bild 2. Gitterfilterstruktur des Synthesefilters



Die zu findende Grundfrequenzperiode wird im Gegensatz zur Autokorrelationsfunktion nicht als Maximum, sondern als Minimum abgebildet. Der entscheidende Vorteil der Differenzfunktion liegt wegen der einfachen Rechnung im Rechenzeitgewinn. Um von der Differenzfunktion zur Grundfrequenzperiode zu gelangen, ist eine Minimumdetektion erforderlich, um die Lage des Minimums zu bestimmen.
Zur stimmhaft/stimmlos-Entscheidung (SH/SL) wird bei den vorliegenden Untersuchungen der normierte Autokorrelationskoeffizient $r\,(1)/r\,(0)$ verwendet, der als besseres Kriterium im Vergleich zur Tiefe des Minimums der Differenzfunktion [6] anzusehen ist.
Als letzte Stufe im Grundfrequenzalgorithmus erfolgt eine Grundfrequenzglättung von jeweils 3 aufeinanderfolgenden Grundfrequenzperioden im Sinne einer Mittelung bei stationären Lautphasen oder einer Korrektur bei den Übergängen stimmhaft/stimmlos oder umgekehrt.
Da der Grundfrequenzalgorithmus u. U. noch Fehlentscheidungen liefern kann, ist es möglich, nach Abschluß der Analyse an Hand des vorliegenden Protokolls Korrekturen am Grundfrequenzverlauf vorzunehmen.

2.2. Syntheseprogramm

Als Syntheseprogramm wurde die Gitterfilterstruktur [4] simuliert. Die Filtergleichungen lauten:

$$\begin{matrix} y_{m-1}\,(n) = y_m\,(n) - k_m\, b_{m-1}\,(n) \\ b_m\,(n) = k_m\, y_{m-1}\,(n-1) + b_{m-1}\,(n-1) \end{matrix}; \quad m = M, \ldots, 7 \qquad (17)$$

Im Bild 2 ist die Gitterfilterstruktur des Synthesefilters dargestellt; Y_{10} ist die Anregung. Entsprechend dem Merkmal SH/SL wird die Anregung zwischen Pitchanregung für stimmhafte und Rauschanregung für stimmlose Laute umgeschaltet. Die Größe der Pitch- wie auch der Rauschanregung ist durch das Merkmal Effektivwert bestimmt. Die Form des Pitchsignals ist halbsinusähnlich. Sie muß entsprechend der Preemphasis bei der Analyse einen $1/\omega$-Frequenzgang haben.
Y_0 ist das synthetisierte Sprachsignal. Es wird zunächst vollständig berechnet, im Hauptspeicher gespeichert und anschließend in Echtzeit über den Analysenmeßplatz ausgegeben.

3. Ergebnisse

Praktische Untersuchungen wurden mit natürlicher Sprache von männlichen und weiblichen Sprechern durchgeführt. Dabei wurden mit dem vorhandenen LPC-Programmsystem gute Ergebnisse bezüglich der Sprachqualität der synthetisierten Sprachproben erzielt.
Wesentliche Voraussetzung für ein fehlerfreies Analyseergebnis ist dabei die Qualität der Tonbandaufzeichnung, die unter Studiobedingungen mit gut artikulierenden Sprechern erfolgen muß.

Eingegangen am 15. September 1982 NaA 8894

Literatur

[1] *Atal, B.; Hanauer, S.:* Speech analysis and synthesis by linear prediction of speech wave. J. Acoust. Soc. Amer. 50 Bd. 2 (1971) 8, S. 637–655
[2] *Makhoul, J.; Wolf, J.:* Linear prediction and the spectral analysis of speech. Bolt, Beranek and Newman, Inc., Boston, Mass., BBN-Rep. 2304 (1972) Aug.
[3] *Makhoul, J.:* Linear Prediction A Tutorial Review. Proc. of the IEEE 63 (1975) 4, S. 561–580
[4] *Markel, J.; Gray, A.:* Linear prediction of speech. Berlin Heidelberg New York: Springer-Verlag 1976
[5] Elektronisches Lernhilfegerät. Patentschrift 2917161. Texas Instruments, Nov. 1979
[6] *Un, C. K.; Yang, S. C.:* A Pitch Extraction Algorithm Based on LPC Inverse Filtering and AMDF. IEEE Trans. ASSP 25 Dez. (1977) 6, S. 565–572

Dr.-Ing. *Gert Jaeger*, VEB Robotron Dresden, ZFT, 8012 Dresden, Postschließfach 330

Perspektiven der Kommunikation mit Rechnern und Automaten

A. Jugel, KDT, Dresden

Mitteilung aus dem VEB Robotron, Zentrum für Forschung und Technik, Dresden

Deskriptoren: Mensch-Maschine-Kommunikation, Stand, Trend; Terminal; Systeme, objekterkennend; Sprachkommunikation; Sprachsynthese, reproduktiv, phonemgesteuert; Spracherkennung

1. Informationsexplosion und Mensch-Maschine-Kommunikation

Das exponentielle Anwachsen der in der Welt produzierten Information wirft die Frage nach Beherrschung dieser Informationsflut auf. Allein mit den menschlichen fünf Sinnen, der Sprache und der Schrift ist der großen Informationsmenge nicht mehr beizukommen.

Dazu ein paar Zahlen und Beispiele:

Der Mensch ist in der Lage, folgende Informationsflüsse aufzunehmen bzw. abzugeben:

Lesen	10	bis 150	bit/s
Hören	10^4	bis 10^5	bit/s
Sehen	10^6	bis 10^8	bit/s
Schreiben	1	bis 50	bit/s
Sprechen	20	bis 150	bit/s.

Dem steht die große Anzahl und Vielfalt der angebotenen Informationen gegenüber. Auf naturwissenschaftlich-technischem Gebiet erscheinen z.Z. jährlich etwa

4000000 Artikel in Fachzeitschriften
300000 neue bzw. neu aufgelegte Bücher
400000 Patentschriften
1000000 Dissertationen und Forschungsberichte.

(Diese Zahlen sind der Durchschnitt, gebildet aus Angaben verschiedener Quellen.)

Diese Informationsmenge verdoppelt sich alle 3 bis 4 Jahre. Eine kleine Rechnung: Ein guter Leser (50 bit/s) beginnt sein Studium 1982 mit dem guten Vorsatz, sich zunächst einen umfassenden Überblick über den Stand von Naturwissenschaft und Technik zu verschaffen. Dazu besorgt er sich die 1982 erschienenen Publikationen (s. Tafel) und beginnt 5 h je Tag zu lesen. Durchschnittlich enthält eine gedruckte Seite 2000 Buchstaben, was rund 10000 bit entspricht. Nach 50 Jahren bei 200 Lesetagen je Jahr hört der Leser auf und stellt fest, daß er gerade ein Hundertstel dessen gelesen hat, was 1982 erschien (Annahme: im Durchschnitt 4 Seiten Artikel, 10 Seiten Patentschrift, 50 Seiten Dissertationen bzw. Forschungsberichte, 100 Seiten Buch). Dabei wurden Prospekte, Produktbeschreibungen, Standards, Gesetzblätter, Tageszeitungen u.v.m. noch nicht einmal berücksichtigt, ganz abgesehen davon, was Rundfunk und Fernsehen an Informationen bieten.

Alles aufzunehmen ist sicher sinnlos und aussichtslos. Man muß auswählen. Damit erscheint ein neues Problem: Wie, was (und zwar das Richtige) finden?

Neben der Informationsaufnahme steht dann das Problem ihrer Verarbeitung und schließlich ihrer Aus- und Weitergabe. Diese Probleme sind nur noch mit moderner Informationstechnik beherrschbar. Kernstück dieser Technik sind die Computer mit ihrer Peripherie zur Kommunikation, Speicherung und Übertragung von Information.

Was diese Technik heute leistet, zeigen folgende Zahlen:

Rechner und ihre Verarbeitungsgeschwindigkeit

— Großrechner	10^7 Operationen/s	$> 10^9$ bit/s
— Kleinrechner	10^6 Operationen/s	$> 10^8$ bit/s
— Mikrorechner	10^6 Operationen/s	$> 10^8$ bit/s

Speicher, Speicherkapazität, Zugriffszeit (für Großrechner)

— Halbleiterspeicher	10^5 bis 10^{10} bit, ms bis ns
— Plattenspeicher	10^9 bit/Stapel, 10 bis 100 ms
— Bandspeicher	10^7 bis 10^9 bit/Band, s bis min
— Bandbibliothek	10^{13} bit, s bis min
— Videoplatten	10^9 bit.

Übertragungsraten

— konventionell bis etwa	10^8 bit/s
— Lichtleiter	$2 \cdot 24 \cdot 10^9$ bit/s (bereits experimentell realisiert).

Betrachtet man nur einmal einen Speicher mit 10^9 bit Speicherkapazität und einer Zugriffszeit von 40 ns/bit, so

- speichert er ≈ 125 Bilder zu je 1024×1024 Bildpunkten und je 64 Farben je Bildpunkt oder den Text von 100000 Buchseiten zu je 2000 Buchstaben (bei 5 bit/Buchstaben)
- ein Bild kann in $\approx 2{,}6$ s auf dem Bildschirm erscheinen
- die 100000 Seiten könnten in 40 s ausgedruckt sein, wenn es einen so schnellen Drucker gäbe (ein schneller Laserdrucker mit 6000 Zahlen/min braucht dazu etwa 14 h). Andererseits würde ein Mensch mit der Schreibmaschine (50 bit/s) als Eingabegerät $\approx 2{,}5$ Arbeitsjahre brauchen, um den Speicher zu füllen.

Mit dieser Gegenüberstellung der menschlichen Kommunikationsfähigkeiten und der Leistung moderner Informationstechnik (die Vergleiche hinken natürlich) sollte folgendes verdeutlicht werden:

Die moderne Informationstechnik kann zwar die Informationsflut auffangen, speichern und auch verarbeiten, beseitigt aber nicht das Problem der Kommunikationsgeschwindigkeit mit dem Menschen. Sie kann nur durch die Anpassung des Rechners an die menschlichen Kommunikationsfähigkeiten erhöht werden.

Damit entsteht ein weiteres Problem:

Mensch und Rechner sprechen unterschiedliche Sprachen. Das wirkt sich besonders hemmend aus, wenn der Rechner angelernt wird, d.h., wenn er seine künstliche Intelligenz durch Programmierung erhält. Zum Beispiel umfaßt das Betriebssystem einer mittleren EDVA $\approx 6 \cdot 10^6$ Anweisungen (Befehle), die in einer der natürlichen Sprache nicht ähnlichen Programmiersprache erstellt werden müssen.

Das bedarf der Arbeit von mehr als 1000 Programmierern über 1 bis 2 Jahre. Dieser Zustand ist insgesamt unbefriedigend. Das heißt, die Schaffung geeigneter Kommunikationsformen ist in Zukunft eine für die Effizienz der Anwendung der Rechentechnik entscheidende Frage. Es geht darum, die Anpassung (Interface) der Maschine mit ihrer „künstlichen Intelligenz" an die Fähigkeiten des Menschen im Sinne eines effektiven Informationsaustausches weiter voranzutreiben und zu neuen Lösungen zu kommen. Das betrifft vor allem

— die Entwicklung neuer, den menschlichen Ausdrucksformen angepaßter Programmiersprachen

— die Weiterentwicklung der „intelligenten" Terminals zur Datenerfassung und Datenmanipulation

— die Entwicklung dialogunterstützender Betriebssysteme sowohl für den durch den Menschen als auch den durch den Rechner gesteuerten Dialog

— die Nutzung der menschlichen Fähigkeit, Befehle zu erkennen, zu verarbeiten und auszugeben durch Entwicklung von Computergrafik- und Bildverarbeitungssystemen (Nutzung des Gesichtssinnes)

— die Entwicklung von Möglichkeiten, in natürlicher Sprache mit dem Rechner zu kommunizieren

— die weitere Reduktion der durch den Menschen zu manipulierenden Daten durch automatische Datenerfassung (z.B. Sensoren, Bild- und Zeichenerkennung) und neue Formen der Datenverarbeitung (z.B. grafische Datenverarbeitung im Rechner) und neue Ausgabeformen (Bilder).

Diese Schwerpunkte zu der weiteren Gestaltung der Mensch-Maschine-Kommunikation entsprechen den technischen Möglichkeiten im Rahmen der Weiterentwicklung der Rechentechnik auf der Basis der Mikroelektronik.
Zur Zeit wird die fünfte Rechnergeneration vorbereitet. Ihre wesentlichen Merkmale werden sein [1] [2]:

— Weitere Erhöhung der Rechengeschwindigkeit und Speicherkapazität durch Einsatz neuer Bauelementegenerationen (z.B. GaAs-Elemente, Josephson-Elemente); damit weitere Miniaturisierung und Verbesserung des Preis-Leistungsverhältnisses

— Vereinfachung der Softwareerstellung (Programmierung) durch neue Softwaretechnologien und Programmiersprachen mit dem Ziel automatischer Programmerstellung und -verifikation. Idealfall: Der Rechner programmiert sich selbst auf der Basis von Vorgaben in natürlicher Sprache

— die Schnittstelle zwischen Mensch und Maschine wird auf der Ebene natürlicher Sprache liegen; (bis 1990 sollen Rechner entstehen, die mit einem Wortschatz von 10000 Wörtern manipulierbar sind).

Alle anderen Formen des Informationsaustausches zwischen Mensch und Rechner, wie alphanumerische Textkommunikation, Bilder und Grafik werden weiterentwickelt:

— Der Rechner soll fähig sein zu lernen, erlerntes Wissen zu speichern und zu verarbeiten, damit wird die assoziative Speicherorganisation die entscheidende Rolle spielen; der Rechner wird mit einer Wissensbank gekoppelt, mit deren Hilfe er Probleme erkennen, verstehen und lösen kann.

Die Vision des sich mit dem Menschen unterhaltenden Computers, der auf Zuruf die notwendigen Informationen bereithält bzw. diese auf Befehl beschafft oder errechnet, ist nicht unreal. Ergebnisse der Grundlagenforschung auf dem Gebiet der Mensch-Maschine-Kommunikation zeigen heute schon die Lösungsmöglichkeiten.

2. Stand und Trend der Mensch-Maschine-Kommunikation

2.1. Intelligente Terminals

Intelligente Terminals sind Nahtstellen zwischen Mensch und Maschine. Ihre Hauptaufgabe ist die gerätetechnische Realisierung des Mensch-Maschine-Dialogs. Das heißt z.B.

— die Übersetzung „menschlicher" Darstellungsformen (Text, Dezimalzahlen, Grafiken, Zeichnungen) in die entsprechende Maschinensprache

— die Steuerung des Dialogs (Bedienerführung) zur effektiven Gestaltung des Dialogs

— die Fehlerüberwachung, -analyse und -beseitigung

— Informationsverarbeitung wie z.B. Normierung, standardgerechte Überarbeitung von Zeichnungen u.a.

Intelligente Terminals sind meist Bildschirmarbeitsplätze mit integriertem Mikrorechner oder Kleinrechner und aufgabenspezifischer Eingabe/Ausgabecharakteristik wie z.B.

— alphanumerische Bildschirme mit alphanumerischer Tastatur (Textverarbeitung, Dialogprogrammierung, Datenerfassung, Laborrechner, Bürocomputer)

— quasigrafische Displays mit alphanumerischer und Zeichentastatur mit Cursor-Steuerung zur Generierung und Korrektur von grafischen Darstellungen

— grafische Displays mit Lichtgriffel, Steuerknüppel, Rollkugel, Tablett oder direkter Eingabe mit Finger (touch screen).

Diese Terminals existieren in großer Vielzahl und sind Stand der Technik. Eine Weiterentwicklung erfolgt z.Z. vor allem auf dem Gebiet grafischer Displays in Richtung hoher Auflösung (4096×4096 Bildpunkte) und großer Anzahl dem Bildpunkt zugeordneter Merkmale (z.B. 256 Farben/Bildpunkt).
Das Hauptproblem dieser Technik besteht in der Software, d.h. in der Art der Intelligenz, mit der sie im Dialog mit dem Menschen arbeiten. Dialogunterstützende Betriebssysteme und dialogfreundliche Sprachen (aufgabenorientiert) sind Hauptgebiet der Forschung und Entwicklung, revolutionierende Fortschritte sind kaum zu erwarten sondern lediglich schrittweise Verbesserungen, solange diese Systeme mit künstlichen (computerorientierten) Sprachen (Programmiersprachen) arbeiten.
Eine neue Qualität entsteht auf dem Gebiet textualer Kommunikation erst beim Übergang zur natürlichen Sprache und ihrer Verarbeitung durch den Computer (s. Abschn. 3.).

2.2. Objekterkennende Systeme

Einen entscheidenden Schritt nach vorn auf dem Gebiet der Datenerfassung sind objekterkennende Systeme. Dazu gehören:

— Belegleser, Zeichenerkenner
— Bilderkennungs- und -verarbeitungssysteme
— Spracherkennungssysteme.

Hier ist die mögliche Erfindungshöhe neuer Entwicklungen noch hoch. Während Belegleser und Zeichenerkenner (begrenzter fest vorgegebener Zeichenvorrat) Stand der Technik sind, stehen Bildverarbeitung und Spracherkennung am Anfang ihrer Entwicklung. Auf diesen Gebieten ist der Mensch mit seinem Nervensystem der Maschine noch weit überlegen.
Mit welch hoher Parallelität und Komplexität die Informationsverarbeitung im Sehsystem vor sich geht, zeigten die Arbeiten von *Hubel* und *Wiesel* (Nobelpreis für Medizin und Physiologie 1981). Erkenntlich ist dabei immer wieder, das sich die hohe funktionelle Komplexität auf wenige Grundprinzipien zurückführen läßt.
Der heutige Stand der Bilderkennung ist durch mehr oder weniger heuristische Verfahren gekennzeichnet, die auf spezielle Anwendungen ausgerichtet sind, wie

— Lageerkennung von Werkstücken

— Erkennen von Bearbeitungsfehlern (z.B. Oberflächenbeschaffenheit)

— Erkennen von definierten Objekten bei Luftaufnahmen usw.

Das Prinzip der Bilderkennung und der Spracherkennung mit heutiger Technik ist etwa gleich

— Digitalisierung des Ausgangssignals (Bild, Wort)

— Analyse des digitalisierten Signals und Merkmalsextraktion (Bilden von Merkmalsvektoren, durch Filterung u.a.)

— Generierung von Referenzvektoren entsprechend dem vorgegebenen Erkennungsbereich (Bildmenge, Wortschatz) auf der Basis von Referenzmodellen der Bilder oder der Sprache

— Vergleich der Merkmalsvektoren mit den Referenzvektoren, Optimierung des Referenzmodells und Generierung eines den Vergleich charakterisierenden Codes (bei Erkennung) zur weiteren Verarbeitung.

Bilderkennung und Sprecherkennung durch informationstechnische Systeme sind Schwerpunkte der Forschung. Am Beispiel der Sprachkommunikation sei das dargestellt.

3. Sprachkommunikation

3.1. Stand, Anwendung, Perspektiven

Sprachkommunikation zwischen Mensch und Maschine in seiner Einheit von

— Spracherkennung — phonetische Eingabe
— Sprachsynthese — phonetische Ausgabe

ist bereits Realität. Sprechendes Spielzeug, sprechende Uhren, Auskunftssysteme, sprachgesteuerte Autos u.v.m. erleben eine erhebliche Umsatzsteigerung. Es folgt die Umsatzentwicklung auf dem Weltmarkt für Spracherkennungs- und -synthesegeräte mit einer durchschnittlichen Steigerungsrate von etwa 425% je Jahr. Keine andere Kommunikationstechnik hat solche Marktchancen.
Das resultiert aus den Vorteilen der Sprachkommunikation und damit der Erschließung neuer Anwendungsgebiete. Wesentliche Vorteile, die aus dem Einsatz von Sprachkommunikationsgeräten resultieren, sind:

— Der Nutzer kann während der Erfüllung von Arbeitsaufgaben, die Hände und Gesichtssinn erfordern, Informationen akustisch aufnehmen bzw. abgeben

— der Nutzer hat eine größere Beweglichkeit am Arbeitsplatz

— die Sprachkommunikation ist unabhängig von den Beleuchtungsverhältnissen

— bei einer Kommandoeingabe kann mit Sprache eine 2- bis 3mal höhere Geschwindigkeit gegenüber Tastaturen erzielt werden; es wird ein höherer Bedienungskomfort erreicht, und die Einarbeitungszeit für ungeübte Nutzer ist kürzer

— für die Übertragung der Sprachinformationen kann ein vorhandenes Telefon- bzw. Sprechfunknetz genutzt werden, so daß die Sprach-E/A-Geräte auch in Fernaufstellung vom Nutzer betrieben werden können.

Nachteile der Sprachkommunikation sind insbesondere die Empfindlichkeit gegenüber Umgebungsgeräuschen sowie die Tatsache, daß die Sprache ein flüchtiger Vorgang ist und somit im Fall einer erforderlichen Archivierung der Originalinformationen zusätzliche Maßnahmen ergriffen werden müssen.
Ein weiterer Nachteil, der gegenwärtig noch Spracherkennungsgeräten anhaftet, ist die starke Sprecherabhängigkeit der Erkennungsresultate, so daß sich jeder neue Sprecher durch einen einmalig durchzuführenden Lernprozeß dem System „bekannt machen" muß (vgl. Abschn. 2.1.), wodurch einer breiten öffentlichen Nutzung gegenwärtig noch Grenzen gesetzt sind. Die genannten Eigenschaften der Sprachkommunikation eröffnen ein breites Anwendungsgebiet von Sprach-E/A-Geräten, insbesondere als Terminals in kommerziellen Auskunftssystemen, für Buchungs- und Lagerhaltungsaufgaben, für Konstrukteur-Arbeitsplätze oder die Programm- und Dateneingabe sowie Korrekturen bei Rechner-Dialogsystemen. Darunter sind sowohl Anwendungsfälle, bei denen Spracheingabe- und -ausgabe gleichzeitig benötigt werden, als auch in anderen Fällen lediglich eine der beiden Funktionen erforderlich ist. Einige konkrete Anwendungsbeispiele sollen das verdeutlichen:

— Rationalisierung der Arbeit an digital-grafischen Arbeitsplätzen („akustische Menütechnik", z.B. bei der Digitalisierung oder Korrektur von Karten, Zeichnungen usw. können Zusatzinformationen, sofern sie mit einem begrenzten Wortschatz beschreibbar sind, akustisch eingegeben werden)

— fortlaufende Speicherung von bestimmten Beobachtungsergebnissen bei speziellen Arbeitsaufgaben (Mikroskopie-Arbeitsplätze für Laborbetrieb, Luftbild-Auswertung, Auswertung von Röntgenaufnahmen bei Reihenuntersuchungen, operative Datenerfassung im Leistungssport u.a.)

— Einsatz der Sprach-Eingabe/Ausgabe in kommerziellen Auskunfts- und Buchungssystemen (z.B. Flugkartenbuchung, Verkehrsauskünfte, Planungs- und Leitungssysteme, Lagerbestandsdateien)

— Steuerung von Sortier- und Registrieranlagen (z.B. Paketsortierung, Qualitätskontrolle, Wareneingangskontrolle, Inventurarbeiten)

— Steuerung und Programmierung von NC-Maschinen und Industrierobotern (Einrichter kann gleichzeitig zu Meß- und Beobachtungsvorgängen Positionieroperationen auslösen)

— Möglichkeit einer Fernbedienung (über drahtloses Mikrofon) bei Bedienerfunktionen in größeren Schaltwerten, Rechenzentren u.ä.; Bediener kann vom jeweiligen Standpunkt aus auf bestimmte (akustisch über Sprachsynthese) angezeigte Zustände reagieren

— Programm- und Dateneingabe für Rechner, insbesondere bei Dialogarbeitsplätzen, Bürocomputereinsatz u.ä.)

— Behinderten-Arbeitsplätze (Ersatz von manuellen Steuerungs- und Eingabemöglichkeiten durch Spracheingabe), z.B. Programmierarbeitsplatz, phonetisch gesteuertes Behindertenauto (1982 auf der Hannover-Messe vorgestellt)

— Einsatz von Sprach-Ausgabeeinrichtungen in Ansage-, Warn- und Alarmanlagen (Schaltwarten, Fahrzeuge, Konsumgüter, militärische Anwendungen); dabei ist ggf. auch eine einfache Übertragung der Informationen über das Telefonnetz möglich. Eine spezielle Anwendung hierbei sind sogenannte Kurzansagegeräte in der Fernmeldetechnik (Zeitansage, Wetterauskunft, ...), die gegenwärtig noch durch wartungsaufwendige mechanische Lösungen (z.B. Magnetband) realisiert werden

— Einsatz von billigen Sprach-Ausgabeeinrichtungen (Schaltkreis-Lösungen) in Massenanwendungen, z.B. „sprechende" Uhr, „sprechender" Taschenrechner, elektronisches Wörterbuch mit Sprachausgabe oder Haushaltsgeräte, heimelektronische Geräte und Fahrzeuge mit akustischer Bediener-Unterstützung

— Anwendungen von Spracherkennern in Konsumgütern (Steuerungsfunktionen an Haushaltsgeräten, Fahrzeugen, Spielzeugen, ...) sind bei dem gegenwärtig erforderlichen technischen Aufwand noch im Anfangsstadium, jedoch mit steigendem Integrationsgrad der Rechner-Schaltkreise bereits absehbar und in großem Umfang durch Einsatz billiger Spezial-Erkennerschaltkreise zu erwarten.

Diese Anwendungsbeispiele zeigen die Perspektive der Sprachkommunikation. Die zu lösenden Aufgaben werden getrennt für Sprachsynthese und Spracherkennung diskutiert.

3.2. Sprachsynthese

Für die technische Realisierung von Sprach-Ausgabegeräten lassen sich zwei prinzipiell unterschiedliche Verfahren nutzen:

a) Reproduktive Synthese, d.h. Wiedergabe speziell komprimierter oder codierter und in ROMs oder EPROMs gespeicherter natürlicher Sprache. Für eine Komprimierung bedient man sich hautpsächlich der aus der Nachrichtentechnik bekannten adaptiven Deltamodulation (ADPCM), womit Datenreduktionen bis auf etwa 10 kbit/s (gegenüber $\approx$ 60 kbit/s bei unkomprimierter Sprache) und eine hohe Natürlichkeit und Verständlichkeit der synthetisierten Sprache erzielt werden können. Als leistungsfähigstes Verfahren für eine Codierung mit speziellen Parametern hat sich das LPC-Verfahren (Linear-Predictive Coding) erwiesen. Damit sind Datenreduktionen bis auf etwa 2 kbit/s erzielbar; jedoch erhöht sich gegenüber der ADPCM der algorithmische Aufwand für die Rückgewinnung des Sprachsignals aus den gespeicherten Parametern auf ein Vielfaches. Als Kernstück eines LPC-Synthetisators ist deshalb stets eine leistungsfähige Multiplizier-einheit (Multiplikationszeit $\approx 5\,\mu s$) erforderlich. Auf Grund der damit verbundenen Kostenunterschiede sowie der mit ADPCM erreichbaren höheren Sprachqualität haben beide Verfahren eine selbständige Daseinsberechtigung; ADPCM bis zu einem gleichzeitig verfügbaren Sprachvolumen von maximal 50 s, LPC bis zu einigen 100 s.
Der Welthöchststand auf dem Gebiet der reproduktiven Synthese wird heute von Halbleiterherstellern, die über leistungsfähige LSI- und VLSI-Technologien verfügen, bestimmt. Je nach Sprachvolumen werden Lösungen, die aus 1 bis 3 Schaltkreisen (darunter 128 kbit-RAMs) bestehen, angeboten.

b) Phonemgesteuerte Synthese bietet im Unterschied zur reproduktiven Synthese, mit der nur ein vom Hersteller generierter Sprachumfang synthetisiert werden kann, auf Grund der Wortbildung aus Phonemen (Lauten) die Möglichkeit zur Synthese eines unbegrenzten Vokabulars. Allerdings muß der Nutzer dem Synthesegerät die Folge der Phonemcodes sowie weitere Steuerinformationen (Tonhöhe, Rhythmus, Akzent) bereitstellen bzw. über ein Programm verfügen, das Buchstabentexte in diese phonetischen Steuerinformationen automatisch umsetzen kann (Graphem-Phonem-Umsetzung).
Die Spitzenleistung für derartige Synthesegeräte stammt von einer USA-Firma, die mit dem Schaltkreis SC-01 einen Baustein herstellt, der aus stimmhaften und stimmlosen Anregungssignalen 63 verschiedene Phoneme künstlich erzeugen und zu fließender Sprache zusammensetzen kann. Auf der Basis dieses Schaltkreises werden drei im Leistungsumfang abgestufte Synthesegeräte hergestellt, das leistungsfähigste mit automatischer Graphem-Phonem-Umsetzung für Englisch als Entwicklungssystem für die Phonemdaten-Generierung.
Bei einer anderen, ebenfalls bedeutsamen Lösung für die phonemgesteuerte Synthese wird, ausgehend von 128 gespeicherten kontext-beeinflußten Lauten (sogenannten Allophonen), die Möglichkeit zur Synthese beliebiger Wörter geschaffen und ebenfalls die Software für die Gewinnung der Allophonfolge aus Buchstabentexten bereitgestellt.
Es läßt sich gegenwärtig noch keine sichere Aussage treffen, welches der beiden vorgestellten Verfahren für die phonemgesteuerte Synthese das zukunftsträchtigere ist. Denkbar ist jedoch, daß der gegenwärtig der phonemgesteuerten Synthese noch anhaftende Nachteil einer geringen Natürlichkeit der synthetischen Sprache durch eine Weiterentwicklung des zweitgenannten Verfahrens leichter zu beseitigen ist. Einen Überblick zu den auf dem Markt befindlichen Geräten zur Sprachsynthese enthält [3].

3.3. Spracherkennung

Ausgehend vom Schwierigkeitsgrad der theoretischen und technischen Beherrschung können folgende 4 Stufen zur Realisierung spracherkennender Systeme charakterisiert werden:

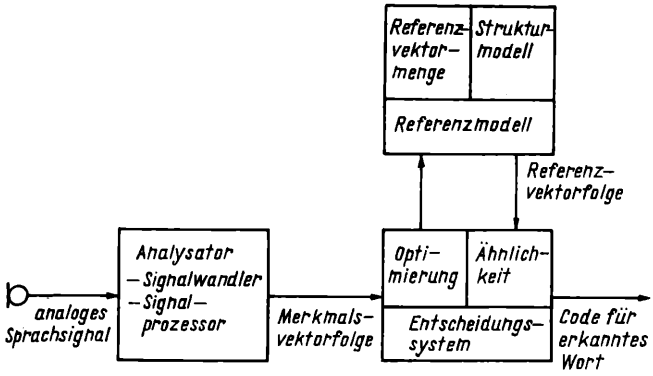


Bild 1. Grundstruktur eines Einzelworterkenners für mittlere Wortzahl

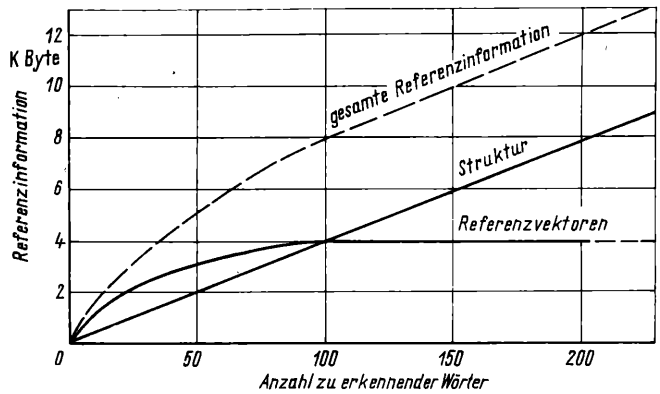


Bild 2. Referenzinformation für Worterkenner in Abhängigkeit von der Wortzahl mit 20 Byte/Referenzvektor, 40 Byte/Wortstruktur

1. Einzelworterkenner für geringen Wortschatz ($\leqq$ 50 Wörter)

2. Einzelworterkenner für mittleren Wortschatz ($\leqq$ 500 Wörter) ggf. mit der Möglichkeit fließender Eingabe kurzer Wortfolgen, die als ein Wort gewertet werden

3. Erkenner für fließend gesprochene Befehle einer künstlichen Sprache mit „einfacher" Syntax und Semantik (einige 1000 Wörter)

4. Erkenner für fließend gesprochene natürliche Sprache ggf. eingeschränkt auf einen begrenzten Kommunikationsbereich (z.B. „Patentanwaltssprache", Fachsprache einzelner Wissenschaftsdisziplinen).

Zusätzlich ist noch zu unterscheiden, ob die Spracherkennung sprecherabhängig ist oder nicht. Bezüglich des Erreichens von Sprecherunabhängigkeit bieten sich folgende Möglichkeiten:

— Anlernen des gesamten Wortschatzes auf jeden neuen Sprecher (sprecherspezifische Referenzinformationen)

— Anpassung der Referenzinformation an einen neuen Sprecher durch Sprechen weniger Codeworte (adaptives Verfahren)

— Benutzung weitgehend sprecherunabhängiger Merkmale, d.h. eines Modells menschlicher Sprache, das sprecherspezifische Besonderheiten eliminiert (Gegenstand der Grundlagenforschung).

Der gegenwärtige Stand ist durch ein breites Angebot von Spracherkennern der Schwierigkeitsstufe 1 und 2 mit Sprecherabhängigkeit gekennzeichnet. Eine Übersicht dazu enthält [3]. Lösungen der Stufe 3 mit zeitweiser Sprecherunabhängigkeit dürften 1985 bis 1990 marktwirksam werden.
Erkenner für fließende Sprache unabhängig vom Sprecher sind nicht vor 1990 zu erwarten. Sie bedürfen noch weiterer Grundlagenforschung, vor allem auf den Gebieten

— linguistischer Modelle der natürlichen Sprache zur Erkennungssteuerung und semantischen Interpretation

— leistungsfähigere Sprachsignalanalysatoren, die eine stärkere Unterstützung sprachlicher Einheiten bei möglichst geringer Sprecherabhängigkeit ermöglichen. Bild 1 zeigt ein Blockdiagramm einer technischen Realisierung eines Einzelworterkenners mit mittlerem Wortschatz.

Das Prinzip besteht darin, aus dem abgetasteten und digitalisierten Sprachsignal durch Signalanalyse (Filterbank) eine Merkmals-Vektorfolge zu erzeugen (Analysator), die mit einer Referenzvektorfolge gebildet im Referenzgenerator verglichen wird. Während des Vergleichsprozesses wird durch Rückkopplung der Suchprozeß im Referenzgenerator optimiert.
Bei Übereinstimmung von Merkmalsvektorfolge und Referenzvektorfolge gibt das Entscheidungssystem einen Code für das erkannte Wort aus, bei 512 Wörtern z.B. minimal 9 bit. Dieser Code wird, dann entsprechend der mit dem Wort verbundenen gewünschten Aktion, durch den nachfolgenden Rechner verwertet (z.B. indem der Rechner das Wort ausdruckt).
Ein solches System erfordert etwa folgenden Hardware-Aufwand:

— Analysator
hochleistungsfähige Signalprozessoren, z.B. µPD 7720 (zusätzlich noch Signalaufbereitung, AD-Wandler usw.)

— Entscheidungssystem
2 Mikroprozessoren der Leistungsfähigkeit U 880 zur Ähnlichkeitsberechnung (1. CPU) und Optimierung und Steuerung der Referenzfolgengenerierung (2. CPU)

— Speicherbedarf
Programmspeicher $\approx$ 8 bis 12 KByte ROM, Arbeitsspeicher $\approx$ 20 bis 28 KByte RAM.

Der Einfluß des Wortschatzes auf die Speichergröße ist dabei gering. Als Referenzinformationen werden z.B. benötigt:

— 20 Byte Referenzvektor, wobei mit 200 Referenzvektoren bereits die Sättigung erreicht wird (ausreichend für beliebige Wortmenge)

— 40 Byte Wort des Vokabulars für die Strukturbeschreibung.

Bild 2 zeigt den benötigten Speicherplatz für die Referenzinformation in Abhängigkeit von der Anzahl der zu erkennenden Wörter. Generelle Probleme der Spracherkennung werden an folgendem Beispiel sichtbar:

— Das Modell für das spracherkennende System geht lediglich von der Struktur des Sprachsignals aus; syntaktische und semantische Aspekte spielen noch keine Rolle

— die zu verarbeitende Informationsmenge zur Beschreibung des Signals und Erkennung des Codes ist so groß, daß damit die Rechenleistung mehrerer Rechner erforderlich wäre, um für fließende Sprache Echtzeiterkennung zu erreichen (z.B. ermöglicht der Einsatz großer EDVA bei 1000 Wörtern und einer beschränkten Anzahl von Texten eine Erkennungsrate von 90%; für die Analyse von 5 min Rede benötigt der Rechner 30 min Rechenzeit [4]).

Eine Ursache für diese Situation liegt in der Struktur der für Spracherkennung eingesetzten Systeme. Sie sind der Struktur des menschlichen Spracherkennungssystems nicht adäquat. Damit ist auch der Prozeß der Erkennung nur unvollkommen nachzubilden. Ein Schwerpunkt weiterer Forschung wird damit (genau wie bei der Bildverarbeitung) die Modellbildung für die natürliche Spracherkennung im Gehirn des Menschen und nicht nur die Modellierung der Sprache selbst sein. Daraus sind Konsequenzen für die Struktur effektiver technischer Erkennungssysteme abzuleiten. Das ist angesichts der großen Bedeutung der Sprachkommunikation in großen Informationssystemen der Zukunft eine dankbare und volkswirtschaftlich wichtige Aufgabe für die Grundlagenforschung.

Eingegangen am 15. September 1982 NaA 8995

Literatur

[1] Angewandte Informatik 24 (1982) 2, S. 158–167
[2] Online (1982) 4, S. 42, 44–48
[3] *Püschel, K.*: Geräte zur phonetischen Eingabe und Ausgabe, gegenwärtiger Stand und Entwicklungstendenzen. Nachrichtentech., Elektron. 33 (1983) 1, S. 18–21
[4] Produktio (1982) 11, S. 3–4
[5] Elektronik 31 (1982) 1, S. 41–42

Dr.-Ing. *Albert Jugel*, VEB Robotron, Zentrum für Forschung und Technik Dresden, 8012 Dresden, Postschließfach 330

Buchbesprechungen

Einführung in die Technologie der Elektrotechnik-Elektronik

Von *H. Eigler* und *W. Beyer*
Berlin: VEB Verlag Technik 1981. 379 Seiten, 195 Bilder, 47 Tafeln, Leinen 26,— M

Das von den Autoren vorgelegte Fachbuch gibt einen guten Überblick über den Umfang der Problematik, so daß der dargestellte Stoff für Fachleute auf technologischem Gebiet und für Fachleute, die technologische Belange beachten müssen, von Interesse ist.
Für den nicht technologisch vorgebildeten Fachmann ist die Abgrenzung zur Technologie des Maschinenbaus, die gesellschaftliche Bedeutung der Technologie und die Kennzeichnung der Technologie als Wissenschaft eine gute Einführung in das Fachgebiet.
Die vorgenommene Systematisierung der technologischen Verfahren nach ihrer naturwissenschaftlichen Erscheinung ermöglicht eine Ordnung der Verfahren, die den logischen Zusammenhang der Teilmengen unterstreicht. Die recht umfangreiche Anzahl der Verfahren wird in einer präzisen, informativen Form behandelt. Verständnis und Vorstellungsvermögen werden durch zahlreiche Bilder wesentlich erleichtert.
Nach der Behandlung der Verfahren wenden sich die Autoren im zweiten Teil des Buches der Fertigungsorganisation zu. In diesem Abschnitt wurden die Gliederung des Produktionsprozesses, die Produktionsorganisation und der Entwurf technologischer Prozesse und Systeme betrachtet. Insbesondere bei der Erläuterung der Entwurfsverfahren ist die instruktive Zusammenstellung verschiedener Modellierungsverfahren hervorzuheben, die einen guten Überblick über die differenzierte Vorgehensweise bei der Gestaltung technologischer Prozesse gibt.
Zusammenfassend kann eingeschätzt werden, daß das vorliegende Buch für Studenten einschlägiger Fachrichtungen, aber auch für Fachleute eine wertvolle Hilfe bei der Einarbeitung in technologische Probleme ist. *H. Preuß* NaB 512

Reallexikon der Akustik

Hrsg. *M. Rieländer*
Frankfurt (Main): Verlag E. Bochinsky 1982, 462 Seiten, 426 Bilder und Tafeln, Kunstleder

Der Verlag Bochinsky, bisher vor allem hervorgetreten mit Publikationen zur musikalischen Akustik, hat mit diesem Lexikon das derzeit wohl umfassendste alphabetische Nachschlagewerk über die gesamte Akustik vorgelegt.
Die Akustik, die Lehre vom Schall, ist ein weitverzweigtes, interdisziplinäres Fachgebiet. Zwar haben Ohrenärzte, Musiker, Werkstoffprüfer, Physiker in der Grundlagenforschung und Ingenieure in der Lärmbekämpfung alle etwas mit Schall zu tun, aber ihre Fachgebiete liegen so weit auseinander, und sie sind so sehr spezialisiert, daß sie Schwierigkeiten bei der Verständigung haben. Deshalb ist ein Lexikon, das die gesamte Akustik umfaßt, uneingeschränkt zu begrüßen. Der behandelte Stoff reicht demzufolge von den physikalischen Grundlagen (Entstehung, Abstrahlung und Ausbreitung von Schall) über Bauakustik (Verminderung von Luft- und Körperschall), Maschinenakustik (Konstruktion geräuscharmer Maschinen), Städtebauakustik, Raumakustik (Herstellung guter Hörsamkeit), Elektroakustik, akustische Meßtechnik bis zur musikalischen Akustik. Selbstverständlich werden auch Schallwahrnehmung und -verarbeitung im Menschen, Schallschädigungen, Spracherkennung und -synthese ebenso behandelt wie etwas am Rande der Akustik liegende Gebiete, z.B. Hydroakustik und Ultraschall.
Es verwundert nicht, daß an diesem umfangreichen Werk 97 Autoren — vorwiegend aus der BRD — gearbeitet haben — ohne Honorar, wie der Herausgeber im Vorwort betont, um den ohnehin schon stattlichen Preis (296,— DM) nicht noch weiter zu erhöhen.
Das Lexikon umfaßt etwa 2000 Stichwörter, wovon ein Teil Verweisstichwörter sind. Die Darstellung ist klar, verständlich und einheitlich; offensichtlich hat der Herausgeber viel Mühe darauf verwendet, die von den Autoren gelieferten Texte aneinander anzupassen. Der Umfang der einzelnen Artikel scheint mir angemessen. So wird das Stichwort „Phasengeschwindigkeit" in 3 Spalten abgehandelt, während für das Stichwort „Normstimmton" 23 Zeilen ausreichen. Die Vollständigkeit eines Lexikons zu beurteilen ist schwierig. Sicherlich wird jeder Fachmann Begriffe vermissen, die er für wichtig hält. Mir fiel auf, daß das moderne Gebiet „Statistische Energieanalyse" nicht enthalten ist, auch die Oberflächenwellenakustik scheint mir etwas kurz weggekommen zu sein. Vorteilhaft ist, daß viele Synonyme aufgenommen sind, auch veraltete. So findet auch der mit der neuesten Terminologie noch nicht vertraute Leser die gesuchte Information und wird gleichzeitig auf bessere oder vorgeschriebene Benennungen hingewiesen. Innerhalb von Artikeln sind Verweisungen nur dann angegeben, wenn in dem zum Verweisstichwort gehörenden Artikel weiterführende Informationen enthalten sind. Die mit Fig. bezeichneten Illustrationen umfassen Schemata, Diagramme, Schaltpläne, Fotos, Notenbeispiele und Tabellen. Viele Artikel sind mit dem Namen des Autors gekennzeichnet; nicht signierte stammen vom Herausgeber oder sind von ihm überarbeitet worden. Unter den meisten Artikeln ist weiterführende Literatur angegeben, und zwar überwiegend moderne bis 1980, in Einzelfällen auch historische. Im Anhang sind die grundlegenden Standardwerke zur Akustik zusammengestellt. Erfreulicherweise gibt der Herausgeber bei vielen Fremdwörtern die Herkunft an z.B. Mikrobar von griech. mikros = klein und baros = Druck. Das verführt ihn aber dazu, durchweg die ph-Schreibweise anzuwenden, was bei vielen Wörtern, bei denen sich die f-Schreibung durchgesetzt hat (Mikrofon, Stereofonie usw.), etwas altertümlich wirkt.
Die Ausstattung läßt keine Wünsche offen; der Satz ist klar und übersichtlich, Druck und Papier sind hervorragend.
R. Fischer NaB 509

Foundations for Microstrip Circuit Design (Grundlagen für den Entwurf von Microstrip-Schaltungen)

Von *T. C. Edwards*
Chichester, New York, Brisbane, Toronto: John Wiley and Sons Ltd., 1981.
265 Seiten, 135 Bilder, 229 Literaturstellen.

Streifenleitungsschaltungen, meist auf der Grundlage der unsymmetrischen Mikrostreifenstruktur, sind seit langem technisch und wirtschaftlich bewährte Bestandteile von Bausteinen, Geräten und Systemen der Hoch- und Höchstfrequenztechnik und damit vor allem der Nachrichtentechnik und Funkortung für zivile und militärische Zwecke. In Mikroelektronik und Rechentechnik sind die charakteristischen Wellenleitereigenschaften dieser Leitungsstruktur um so bedeutsamer, je kürzer Impulsanstiegszeit und Impulsdauer werden.
Im vorliegenden Buch gibt der Autor eine deutlich auf die praktische Anwendung orientierte Zusammenstellung der für die Analyse und auch Synthese solcher Leitungsbausteine und -schaltungen erforderlichen Grundlagen. Nach einer kurzen Wiederholung der Grundbegriffe aus der Leitungstheorie für TEM- oder Lecherleitungen und einem Überblick über die verschiedenen Leitungsstrukturen, die für integrierte Mikrowellenschaltungen prinzipiell in Betracht kommen, sowie über die zugehörigen Werkstoffe und Technologien werden in 2 Kapiteln die Leitungskenngrößen und die Algorithmen zur Dimensionierung von Microstrip-Leitungen behandelt.
Unter Bezugnahme auf die feldtheoretischen Untersuchungen, wie sie von *Wheeler* und anderen Autoren dargelegt wurden, werden hier die daraus ableitbaren Näherungsbeziehungen für die bequeme rechnergestützte Analyse und Synthese von solchen Leitungen zusammengestellt und hinsichtlich Gültigkeitsbereich und Genauigkeit kritisch miteinander verglichen. Im Vordergrund stehen dabei die in der Praxis interessierenden Wertebereiche. Das gilt sowohl für die Betrachtung der Kennwerte von Leitungen mit Quasi-TEM-Welle bei nicht zu hohen Frequenzen, für die Berücksichtigung der Dispersionserscheinungen und weitere Wellentypen bei höheren Frequenzen als auch für die Behandlung der Schaltungskennwerte der verschiedensten Leitungsdiskontinuitäten (Kap. 5) und verkoppelter Streifen-

leitungen und Richtkoppler (Kap. 6) und ihrer Frequenzabhängigkeit. Übergänge von anderen Leitungsstrukturen auf die Microstrip-Leitung und die Besonderheiten der Meßtechnik für die Messung von Substrat- und Leitungskennwerten werden in einem weiteren Kapitel erörtert. Die Anwendung der bisher dargestellten Entwurfsrichtlinien wird an zehn repräsentativen Microstrip-Schaltungen (Filter, Verstärker u.a.) demonstriert.
Sehr zu begrüßen sind die deutlich formulierten Empfehlungen am Ende mehrerer Kapitel, die ausdrücklichen Hinweise und Strategien für die Anwendung moderner Rechentechnik (besonders auch in Form von Taschen- und Tischrechnern), die prägnante Darstellung des gesamten Stoffs und die ausführlichen Literaturangaben. Das Buch ist für die in der Praxis tätigen Entwicklungsingenieure der Hoch- und Höchstfrequenztechnik ein nützliches anwendungsorientiertes Arbeitsmittel.

M. Kummer NaB 500

Electronic Noice (Elektronisches Rauschen)

Von *A. Ambrozy*.
Budapest: Akademiai Kiado 1982. 281 Seiten, 143 Bilder, 18 Aufgaben, 544 Literaturquellen, 62,50 M

Das aus dem Ungarischen übersetzte und gemeinsam mit dem Verlag McGraw-Hill, New York, in englischer Sprache herausgegebene Werk ist ein vorzügliches Einführungswerk zur Gesamtproblematik des elektronischen Rauschens. Dieses physikalische Phänomen hat in der gesamten Informations- und Automatisierungstechnik entscheidende Bedeutung erlangt, weil wir durch Kenntnis der zugehörigen Gesetze z.B. Auskunft über die natürliche Grenze für die Übertragung schwacher Signale oder über die erzielbare Genauigkeit einer Messung erhalten. Da Rauscherscheinungen praktisch in allen elektronischen Bausteinen und Einrichtungen präsent sind, hat diese „Fehlerquelle" — das ist meist der dominierende Aspekt, wenn wir von Anwendungen in der Meßtechnik absehen — Querschnittscharakter und ist gleichermaßen für den Natur- und Ingenieurwissenschaftler von Gewicht.
Inhaltlich behandelt der Autor, der als Hochschullehrer an der Technischen Universität Budapest tätig ist, nach einer schon mit Basisgesetzen angereicherten Einleitung in weiteren sieben Abschnitten

— Verteilungs- und Dichtefunktionen
— Stochastische Signale im Zeit- und Frequenzbereich
— Physikalische Aspekte des Rauschens
— Rauschparameter linearer Netzwerke
— Rauschen in Halbleiterschaltungen
— Effekte bei nichtlinearen Übertragungseigenschaften und
— Probleme bei der Messung von Rauschkenngrößen.

Die hinsichtlich Text, Formeln und Bildern ausgewogene Darstellung wird nach jedem Abschnitt durch ausgewählte, z.T. sehr originelle Aufgaben (mit Lösungen) und durch ausführliche Literaturlisten ergänzt. Die klare, gut lesbare und stets verständliche, textliche Formulierung erleichtert das Studium des Werkes, obgleich an zahlreichen Einzelstellen auf die Beweisführung — und damit auf die Voraussetzungen und den Gültigkeitsbereich — für die zitierten Gesetze verzichtet wird. Für den Lernenden wäre es zusätzlich vorteilhaft, wenn die mitgeteilten Erkenntnisse eine Wichtung bezüglich ihrer Praxiswirksamkeit erhalten hätten. Vom pädagogischen Standpunkt wird es als nicht glücklich eingeschätzt, daß für Frequenzkennfunktionen nebeneinander Groß- und Kleinbuchstaben — z.B. zur Kennzeichnung der Leistungsdichte im physikalischen und im systemtheoretischen Sinne — eingeführt werden. Hier sollte konsequent an der bewährten Regel: Zeitfunktionen mit kleinen und Frequenzfunktionen mit großen Buchstaben beschreiben, festgehalten werden.
Alles in allem wird sowohl den Studenten der physikalischen und elektrotechnischen Fakultäten als auch in der Praxis tätigen Physikern und Elektroingenieuren ein Titel präsentiert, der hält, was er verspricht. Durch die zahlreichen Literaturstellen (zitiert wird bis 1978) findet der an tieferen Zusammenhängen interessierte Nutzer bequem die notwendige Hintergrund- und Quelleninformation. *G. Fritzsche* NaB 515

TECHNIK-Literatur

Bestellschein

Die aufgeführten Bücher können Sie mit diesem Bestellschein im Inland beim **örtlichen Buchhandel** bestellen.

Gerätekonstruktion

Autorenkollektiv

Federführung und Gesamtredaktion: Prof. Dr.-Ing. habil. *Werner Krause*

1. Auflage, 660 Seiten, 501 Bilder, 219 Tafeln
Ganzleinen 49,— M
Bestell-Nr.: 553 088 1 _____

Stahlbau — Band 1

Berechnung und Bemessung der Elemente von Stahlkonstruktionen

Prof. em. Dr.-Ing. Dr. techn. h.c. *Friedrich Kurth*

11., bearbeitete Auflage, 222 Bilder, 97 Tafeln
Pappband 9,80 M
Bestell-Nr.: 553 041 9 _____

Einführung in die Elektrotechnik — Arbeitsbuch

Prof. Dr.-Ing. habil. *Klaus Lunze* und Dr.-Ing. *Eberhart Wagner*

4., bearbeitete Auflage, 296 Seiten
Kunstleder 20,— M
Bestell-Nr.: 553 072 6 _____

TECHNIK-WÖRTERBUCH

Elektrotechnik — Elektronik
deutsch — englisch
mit etwa 60000 Wortstellen

Herausgegeben von Prof. Dr. sc. techn. *Peter-Klaus Budig*

1. Auflage, 700 Seiten
Kunstleder 55,— M
Bestell-Nr.: 553 113 9 _____

Darüber hinaus können Sie alle in diesem Heft angezeigten und rezensierten Bücher des VEB Verlag Technik mit diesem Vordruck bestellen.

_____ _____

_____ _____

_____ _____

_____

Name, Vorname Anschrift mit Postleitzahl

_____

Datum Unterschrift

SOMMAIRE

16e Colloque sur l'Informatique 1983.
Acoustique et Technique de Mesure
2e page de couverture

Hart, H.
La technique de mesure en évolution. L'influence de la micro-électronique sur la technique de mesure
p. 4 – 8

Krambeer, H.
L'essai de circuits analogiques avec des sources contrôlées
p. 8 – 10

Seifart, M.
Un système bifilaire de collecte de données de mesure avec numérisation près des senseurs
p. 10 – 11

Brovkov, W.
Un convertisseur analogique-numérique à 12 bits avec traitement parallèle échelonné
p. 12 – 14

Kranz, E. G.
Algorithmes pour la détermination numérique de la valeur effective
p. 14 – 16

On peut déterminer la valeur effective de signaux sous la forme de la valeur moyenne temporelle ou de la valeur moyenne d'ensemble. On présente les algorithmes correspondants pour le traitement de valeurs d'échantillonage et on les compare en ce qui concerne leur effectivité.

Uhlig, G.
La génération numérique de fonctions selon la méthode tabulaire
p. 16 – 18

La technique automatique d'essai et de mesure dépend essentiellement de la génération programmable de signaux. Les principes de génération numérique de signaux, réalisés à l'aide d'un microcalculateur, sont capables de délivrer des signaux fixes dans une large gamme de paramètres.

Püschel, K.
Dispositifs d'entrée et de sortie phonétiques. Etat actuel et tendances du développement
pp. 21 – 26

Blutner, F.
Modèle fonctionnel psychoacoustique et possibilités d'une optimalisation de systèmes techniques
p. 21 – 26

On présente le projet d'un système effectif de sélection de caractéristiques acoustiques qui est basé sur un modèle psychoacoustique. On se sert de quelques exemples d'application d'un tel processeur acoustique flexible et universel pour étudier des problèmes du traitement de valeurs de mesure de l'acoustique vocale et musicale.

Kordon, U.
La méthode de l'entassement pour l'évaluation de méthodes d'analyse dans le traitement vocal
p. 26 – 28

Affa, K.-H.
Une méthode pour l'adaptation dynamique des speakers
p. 28 – 30

Weigert, G.
Un banc de filtres numériques pour l'analyse vocale à l'aide du KRS 4201
p. 30 – 32

On décrit le principe de fonctionnement d'un banc parallèle de filtres numériques. A l'aide d'un programme établi pour le petit ordinateur KRS 4201, on réalise la simulation d'un banc de filtres à 18 voies ce qui permet d'analyser des signaux vocaux.

Haller, K. et *Blutner, F.*
Un modèle multiple pour la description d'attributs sonores subjectifs
p. 32 – 35

Jaeger, G.
La collecte de données vocales selon le principe du codage à prédiction linéaire
p. 36 – 37

On explique le fonctionnement du codage à prédiction linéaire pour la compression de données vocales. On décrit un système de programmes qui permet la compression et la synthèse de la voix humaine à l'aide de ce principe. L'analyse effectue le calcul du coefficient CPL et de la fréquence fondamentale du signal vocal.

Jugel, A.
Perspectives de la communication avec des ordinateurs et des automates
p. 38 – 41

Nachrichtentechnik · Elektronik

Herausgeber	Kammer der Technik
Verlag	VEB Verlag Technik DDR — 1020 Berlin, Oranienburger Str. 13/14 Telegrammadresse: Technikverlag Berlin Telefon: 28700; Telex: 0112228 techn dd
Verlagsdirektor	Dipl. oec. *Herbert Sandig*
Redaktion	Ing. *Günther Barth*, Verantwortlicher Redakteur (Telefon: 2870332); Dipl.-Ing. *Christa Drawert*, Redakteur (Telefon: 2870348) Ing. *Gisela Rehahn*, Redakteur (Telefon: 2870347);
Lizenz-Nr.	1103 des Presseamtes beim Vorsitzenden des Ministerrates der Deutschen Demokratischen Republik
AN (EDV)	18432
Erscheinungsweise	monatlich 1 Heft
Heftpreis	4,– M, Abonnementpreis vierteljährlich 12,– M; Auslandspreise sind den Zeitschriftenkatalogen des Außenhandelsbetriebes BUCHEXPORT zu entnehmen.
Gesamtherstellung	Messedruck Leipzig, BT Borsdorf, 7122 Borsdorf III-18-328
Anzeigenannahme	Für Bevölkerungsanzeigen alle Anzeigen-Annahmestellen in der DDR, für Wirtschaftsanzeigen der VEB Verlag Technik, 1020 Berlin, Oranienburger Str. 13/14, PSF 201. Anzeigenpreisliste Nr. 6; Auslandsanzeigen: Interwerbung GmbH, DDR — 1157 Berlin, Hermann-Duncker-Str. 89
Erfüllungsort und Gerichtsstand	Berlin-Mitte. Der Verlag behält sich alle Rechte an den von ihm veröffentlichten Aufsätzen und Abbildungen, auch das der Übersetzung in fremde Sprachen, vor. Auszüge, Referate und Besprechungen sind nur mit voller Quellenangabe zulässig.
Bezugsmöglichkeiten	
DDR	sämtliche Postämter
SVR Albanien	Direktorije Quendrore e Perhapjes dhe Propaganditit te Librit Rruga Konference e Pezes, Tirana
VR Bulgarien	Direkzia R.E.P., 11a, Rue Paris, Sofia
VR China	China National Publications Import and Export Corporation, West Europe Department, P.O. Box 88, Beijing
ČSSR	PNS — Ustredni Expedicia a Dovoz Tisku Praha, Vinohradská 41, 12505 Praha PNS, Ustred na Expedicia Tlače, Gottwaldovo nám. 48, 88419 Bratislava
SFR Jugoslawien	Jugoslovenska Knjiga, Terazije 27, Beograd; Izdavačko Knjižarsko Produzeće MLADOST, Ilica 30, Zagreb
Koreanische DVR	CHULPANMUL Korea Publications Export & Import Corporation, Pyongyang
Republik Kuba	Empresa de Comercio Exterior de Publicaciones, O'Reilly No. 407, Ciudad Habana
VR Polen	C. K. P. i W. Ruch, Towarowa 28, 00-958 Warszawa
SR Rumänien	Directia Generala a Postei si Difuzarii Presei, Palatul Administrativ, Bucuresti
UdSSR	Städtische Abteilungen von Sojuzpečat' oder Postämter und Postkontore
Ungarische VR	P. K. H. I., Külföldi Elöfizetési Osztály, P.O. Box 16, 1426 Budapest
SR Vietnam	XUNHASABA, 32, Hai Ba Trung, Hanoi
BRD und Berlin (West)	Brücken-Verlag GmbH, Ackerstraße 3, 4000 Düsseldorf 1; ESKABE Kommissions-Grossobuchhandlung, Postfach 36, 8222 Ruhpolding/Obb.; Helios Literatur-Vertriebs-GmbH, Eichborndamm 141 – 167, Berlin (West) 52, Kunst und Wissen Erich Bieber OHG, Postfach 46, 7000 Stuttgart 1; Gebrüder Petermann, BUCH + ZEITUNG INTERNATIONAL, Kurfürstenstraße 111, Berlin (West) 30
Österreich	Helios Literatur-Vertriebs GmbH & Co. KG, Industriestraße B 13, A-2345 Brunn am Gebirge
Schweiz	Genossenschaft Literaturvertrieb, Cramerstr. 2, 8004 Zürich
Alle anderen Länder	örtlicher Buchhandel; BUCHEXPORT Volkeigener Außenhandelsbetrieb der Deutschen Demokratischen Republik, DDR — 7010 Leipzig, Postfach 160; und Leipzig Book Service, DDR — 7010 Leipzig, Talstraße 29